



# Multiple knockout mouse models reveal lincRNAs are required for life and brain development

## Citation

Sauvageau, M., L. A. Goff, S. Lodato, B. Bonev, A. F. Groff, C. Gerhardinger, D. B. Sanchez-Gomez, et al. 2013. "Multiple knockout mouse models reveal lincRNAs are required for life and brain development." *eLife* 2 (1): e01749. doi:10.7554/eLife.01749. <http://dx.doi.org/10.7554/eLife.01749>.

## Published Version

doi:10.7554/eLife.01749

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:11879469>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# Multiple knockout mouse models reveal lincRNAs are required for life and brain development

Martin Sauvageau<sup>1,2†</sup>, Loyal A Goff<sup>1,2,3†</sup>, Simona Lodato<sup>1,2†</sup>, Boyan Bonev<sup>1,2</sup>, Abigail F Groff<sup>1,2</sup>, Chiara Gerhardinger<sup>1,2</sup>, Diana B Sanchez-Gomez<sup>1</sup>, Ezgi Hacisuleyman<sup>1,2</sup>, Eric Li<sup>1</sup>, Matthew Spence<sup>1</sup>, Stephen C Liapis<sup>1,2</sup>, William Mallard<sup>1,2</sup>, Michael Morse<sup>1,2</sup>, Mavis R Swerdel<sup>4</sup>, Michael F D'Ecclesiss<sup>4</sup>, Jennifer C Moore<sup>5</sup>, Venus Lai<sup>6</sup>, Guochun Gong<sup>6</sup>, George D Yancopoulos<sup>6</sup>, David Friendewey<sup>6</sup>, Manolis Kellis<sup>2,3</sup>, Ronald P Hart<sup>4</sup>, David M Valenzuela<sup>6</sup>, Paola Arlotta<sup>1,2\*</sup>, John L Rinn<sup>1,2,7\*</sup>

<sup>1</sup>Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, United States; <sup>2</sup>Broad Institute of MIT and Harvard, Cambridge, United States; <sup>3</sup>Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, United States; <sup>4</sup>Department of Cell Biology and Neuroscience, Rutgers, The State University of New Jersey, New Brunswick, United States; <sup>5</sup>Department of Genetics, Rutgers, The State University of New Jersey, New Brunswick, United States; <sup>6</sup>VelociGene, Regeneron Pharmaceuticals Inc., Tarrytown, United States; <sup>7</sup>Department of Pathology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, United States

**\*For correspondence:**  
Paola\_Arlotta@harvard.edu (PA);  
johnrinn@fas.harvard.edu (JLR)

<sup>†</sup>These authors contributed  
equally to this work

**Competing interests:** See page 20


**Funding:** See page 20

**Received:** 21 October 2013

**Accepted:** 21 November 2013

**Published:** 31 December 2013

**Reviewing editor:** Danny  
Reinberg, Howard Hughes  
Medical Institute, New York  
University School of Medicine,  
United States

 Copyright Sauvageau et al.  
This article is distributed under  
the terms of the [Creative  
Commons Attribution License](#),  
which permits unrestricted use  
and redistribution provided that  
the original author and source  
are credited.

**Abstract** Many studies are uncovering functional roles for long noncoding RNAs (lncRNAs), yet few have been tested for in vivo relevance through genetic ablation in animal models. To investigate the functional relevance of lncRNAs in various physiological conditions, we have developed a collection of 18 lncRNA knockout strains in which the locus is maintained transcriptionally active. Initial characterization revealed peri- and postnatal lethal phenotypes in three mutant strains (*Fendrr*, *Peril*, and *Mdgt*), the latter two exhibiting incomplete penetrance and growth defects in survivors. We also report growth defects for two additional mutant strains (*linc-Brn1b* and *linc-Pint*). Further analysis revealed defects in lung, gastrointestinal tract, and heart in *Fendrr*<sup>-/-</sup> neonates, whereas *linc-Brn1b*<sup>-/-</sup> mutants displayed distinct abnormalities in the generation of upper layer II–IV neurons in the neocortex. This study demonstrates that lncRNAs play critical roles in vivo and provides a framework and impetus for future larger-scale functional investigation into the roles of lncRNA molecules.

DOI: 10.7554/eLife.01749.001

## Introduction

Mammalian genomes encode thousands of long noncoding RNAs (lncRNAs), which are emerging as key regulators of cellular processes (Rinn and Chang, 2012; Mercer and Mattick, 2013). Gain- and loss-of-function approaches in cell-based in vitro systems have been useful in uncovering important roles for lncRNAs, such as modulating chromatin states, maintaining cellular identity (i.e. pluripotency) and regulating cell cycle and translation (Tsai et al., 2010; Guttman et al., 2011; Wang et al., 2011; Yoon et al., 2012). Genome-wide association and expression profiling studies in humans have also identified correlations between lncRNA mutation, misregulation and disease states (Visel et al., 2010; Cabili et al., 2011; Brunner et al., 2012). Yet, direct in vivo, genetic evidence of the functional significance of lncRNAs as a class of transcripts remains elusive.

**eLife digest** The mammalian genome is comprised of DNA sequences that contain the templates for proteins, and other DNA sequences that do not code for proteins. The coding DNA sequences are transcribed to make messenger RNA molecules, which are then translated to make proteins. Researchers have known for many years that some of the noncoding DNA sequences are also transcribed to make other types of RNA molecules, such as transfer and ribosomal RNA. However, the true breadth and diversity of the roles played by these other RNA molecules have only recently begun to be fully appreciated.

Mammalian genomes contain thousands of noncoding DNA sequences that are transcribed. Recent *in vitro* studies suggest that the resulting long noncoding RNA molecules can act as regulators of transcription, translation, and cell cycle. *In vitro* studies also suggest that these long noncoding RNA molecules may play a role in mammalian development and disease. Yet few *in vivo* studies have been performed to support or confirm such hypotheses.

Now Sauvageau et al. have developed several lines of knockout mice to investigate a subset of noncoding RNA molecules known as long intergenic noncoding RNAs (lincRNAs). These experiments reveal that lincRNAs have a strong influence on the overall viability of mice, and also on a number of developmental processes, including the development of lungs and the cerebral cortex.

Given that the vast majority of the human genome is transcribed, the mouse models developed by Sauvageau et al. represent an important step in determining the physiological relevance, on a genetic level, of the noncoding portion of the genome *in vivo*.

DOI: [10.7554/eLife.01749.002](https://doi.org/10.7554/eLife.01749.002)

Functional studies in knockout mouse models have provided compelling evidence for the requirement and sufficiency of particular transcripts for organ development and function. However, small and large-scale efforts have focused primarily on protein coding genes, leaving long noncoding RNAs vastly understudied (Edwards et al., 2011; White et al., 2013). Of the few lincRNA mutant mice generated to date, most involved previously studied and classic examples of lincRNAs (e.g., *Xist*, *H19*, *Kcn11ot1*, *Malat1*) (Anguera et al., 2011; Gomez et al., 2013; Gordon et al., 2010; Moseley et al., 2006; Nakagawa et al., 2011; Ripoché et al., 1997; Zhang et al., 2012), leaving the more recent large-scale RNA-Seq-derived catalogs of lincRNAs largely unexplored. The progress made in these early studies to initiate functional characterization of lincRNAs *in vivo* now warrants a larger investigation of their biological importance to development and disease.

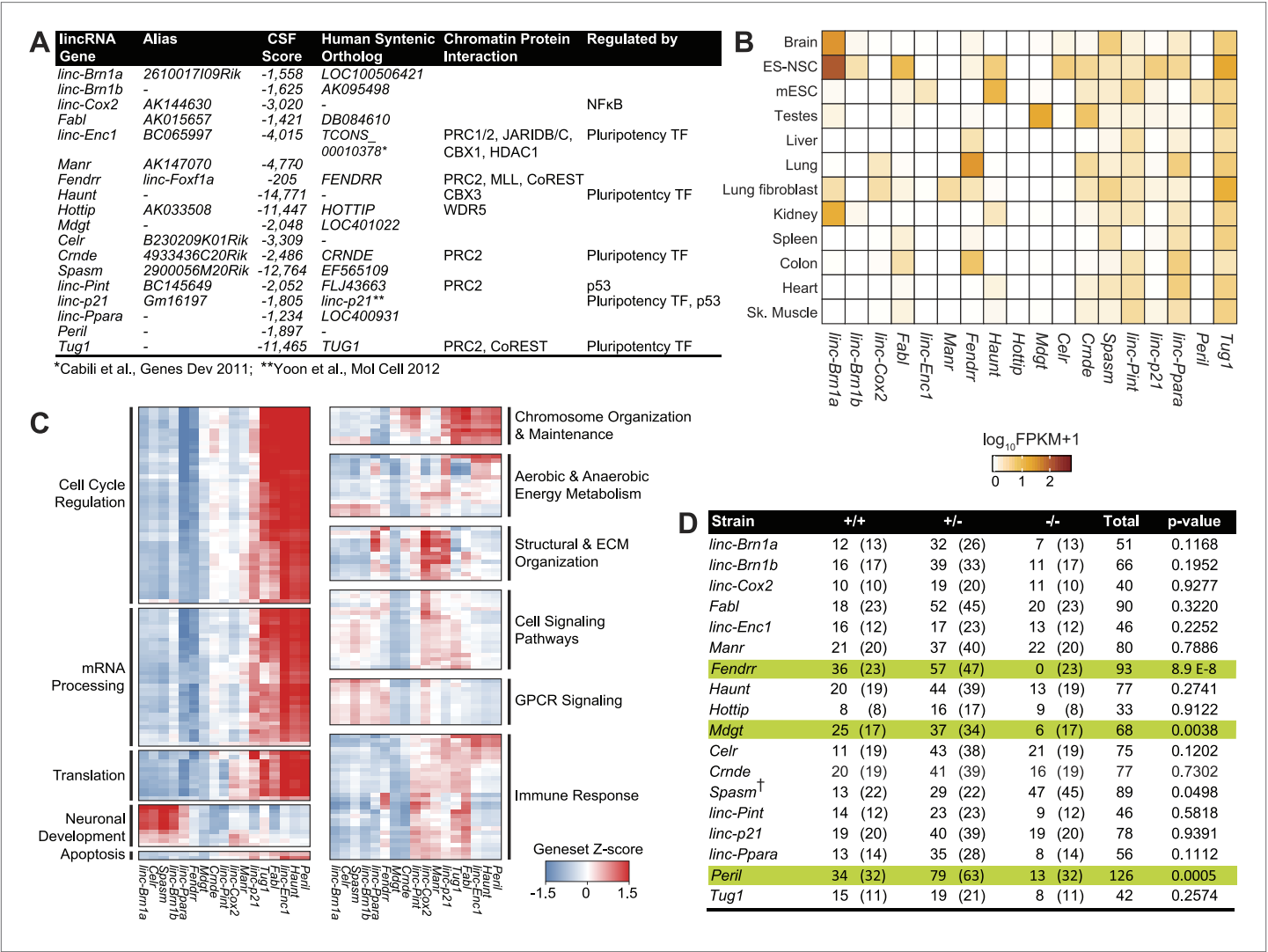
To address this directly, and to begin to explore the roles of lincRNAs *in vivo*, we have established a novel cohort of lincRNA knockout mice. We focused on a subgroup of lincRNAs called long intergenic noncoding RNAs (lincRNAs), such that genetic deletion would not overlap known protein coding genes or other known gene annotations. We have implemented a generalized and logical lincRNA candidate selection process that leverages a collection of cell-based functional assays, RNA-sequencing data and computational analyses. This approach led us to identify 18 lincRNAs for targeted deletion in mouse. Initial characterization of these new knockout strains demonstrated key functional roles in viability, development of the cerebral cortex and other developmental processes. In this study, we describe the observed phenotypes for five strains within this collection. Collectively, these data provide evidence that lincRNAs play central roles in mammalian development and physiology.

## Results

### Global characterization of lincRNA knockout candidate selection

To more globally understand the physiological significance of lincRNAs, we combined a step-wise lincRNA selection pipeline with a genetic approach to engineer a cohort of lincRNA knockout mouse strains for functional analysis. We integrated several computational and experimental data sets to select *bona fide* lincRNA candidates across three lincRNA catalogs (Guttman et al., 2009; UCSC and Refseq). First, all transcripts with identifiable Pfam domains or those overlapping known non-lincRNA annotations (e.g., annotated protein coding genes, microRNAs, tRNAs and pseudogenes) were excluded. Second, we excluded any remaining transcripts with conserved protein coding potential, including

the potential to produce small peptides, by performing stringent codon-substitution frequency (CSF) analysis (Lin et al., 2011). We previously demonstrated that this algorithm is capable of discriminating known functional small peptides down to 11 amino acids (Guttman and Rinn, 2012; Guttman et al., 2013). Using selective criteria, we restricted candidate lincRNAs to those with CSF scores <−200 (ranging from −205 to −14,771, Figure 1A). Then, we examined existing ribosome profiling datasets to quantitate the ability of these transcripts to engage the ribosome (Figure 1—figure supplement 1). None of the tested candidates were shown to have clear translation efficiency or codon bias according



**Figure 1.** Properties of the 18 lincRNA candidates and Mendelian inheritance. **(A)** List of the 18 lincRNA candidates for targeted deletion in mouse and overview of criteria used for their selection. **(B)** Heatmap of log<sub>10</sub> FPKM+1 expression levels of the 18 lincRNAs in a panel of adult mouse tissues and cell lines via RNA-Seq. **(C)** Guilt-by-association (GBA) analysis for 17/18 lincRNA candidates. Individual tiles represent significant ( $p < 1.0 \times 10^{-6}$ ) gene sets from the CP Reactome collection at MSigDB. Tiles are filled based on the Z-score of the Pearson correlation values between a given lincRNA and the genes within the gene set across a compendium of RNA-Seq samples. **(D)** Mendelian inheritance of the 18 lincRNA mutant alleles from the progeny of heterozygote intercrosses. Numbers of observed and expected (in parenthesis) wild-type (+/+), heterozygote (+/-) and homozygote (-/-) mice are indicated. Mice were genotyped at weaning age. The p value is based on X2 test. † The *Spasm* gene is located on the X chromosome.

DOI: 10.7554/eLife.01749.003

The following figure supplements are available for figure 1:

**Figure supplement 1.** Ribosome release score of lncRNAs.

DOI: 10.7554/eLife.01749.004

**Figure supplement 2.** lincRNA candidates knockout targeting strategy.

DOI: 10.7554/eLife.01749.005

to proposed standards (Guttman et al., 2013). Finally, through genome lift-over of our mouse lincRNAs to the human genome (build hg19), we examined mass spectrometry data by Gascoigne et al. (2012) to discard transcripts with mapped peptides. We kept transcripts with only two or less mass spectrometry tags, consistent with low levels of background ribosome association observed by Guttman et al. (2013) (Supplementary file 1B). Thus, based on an exhaustive analysis of existing annotations, standard evolutionary and ribosome profiling metrics, as well as mass spectrometry data, these candidates do not appear to contain protein coding potential.

From this stringent list of lincRNAs, we selected 18 candidates by further considering additional criteria: (1) Presence of an observed orthologous human lincRNA with syntenic conservation as determined by TransMap (Zhu et al., 2007); (2) Presence of canonical chromatin features of actively transcribed genes such as histone H3 lysine 4 trimethylation (H3K4me3) at the promoter and histone H3 lysine 36 trimethylation throughout the gene body (Guttman et al., 2009, 2010); (3) Local 'enhancer-like' signatures based on published P300, RNA polymerase II, histone lysine 4 monomethylation (H3K4me1) and histone lysine 27 acetylation (H3K27ac) chromatin states (Shen et al., 2012). We selected 6/18 candidate lincRNA loci that have putative 'enhancer' modifications, thus allowing to investigate the newfound roles of lincRNAs in 'enhancer' activity (Ørom et al., 2010; Wang et al., 2011). Collectively, this step-wise selection process resulted in 18 high-quality candidate lincRNAs with diverse features, for targeted deletion and phenotypic characterization (Figure 1A, Supplementary file 1A for coordinates and nomenclature details).

We first examined the gene-expression patterns of the candidate lincRNAs using RNA-sequencing of various adult tissues and cell types. Several lincRNAs presented more restricted patterns of expression, suggesting strong tissue specificity, for example *Celr* in embryonic stem (ES)-derived neural stem cells (NSCs), *linc-Enc1* in mouse ES cells, *Manr* and *linc-Cox2* in lung, whereas a select few showed more ubiquitous expression across tissues (*linc-Pint*, *Spasm*, *linc-Ppara*, *Tug1*) (Figure 1B). Most of the candidate lincRNAs (12/18) were expressed in the adult brain or in ES-derived NSCs (Figure 1B).

Since most candidate lincRNAs in this screen have unknown biological roles, we leveraged our RNA-sequencing data using guilt-by-association (GBA) analysis to generate hypotheses on functional significance by comparing the expression of each lincRNAs to protein coding genes of known function ('Materials and methods') (Figure 1C). GBA predicted lincRNA activities across a wide range of pathways and biological processes, ranging from regulation of the cell cycle and chromosome organization and maintenance, to neuronal differentiation and immune response (Figure 1C). Overall, these analyses suggested diverse potential roles for our candidate lincRNAs.

## Genetic analysis of lincRNA knockout mice reveals key roles in mammalian development

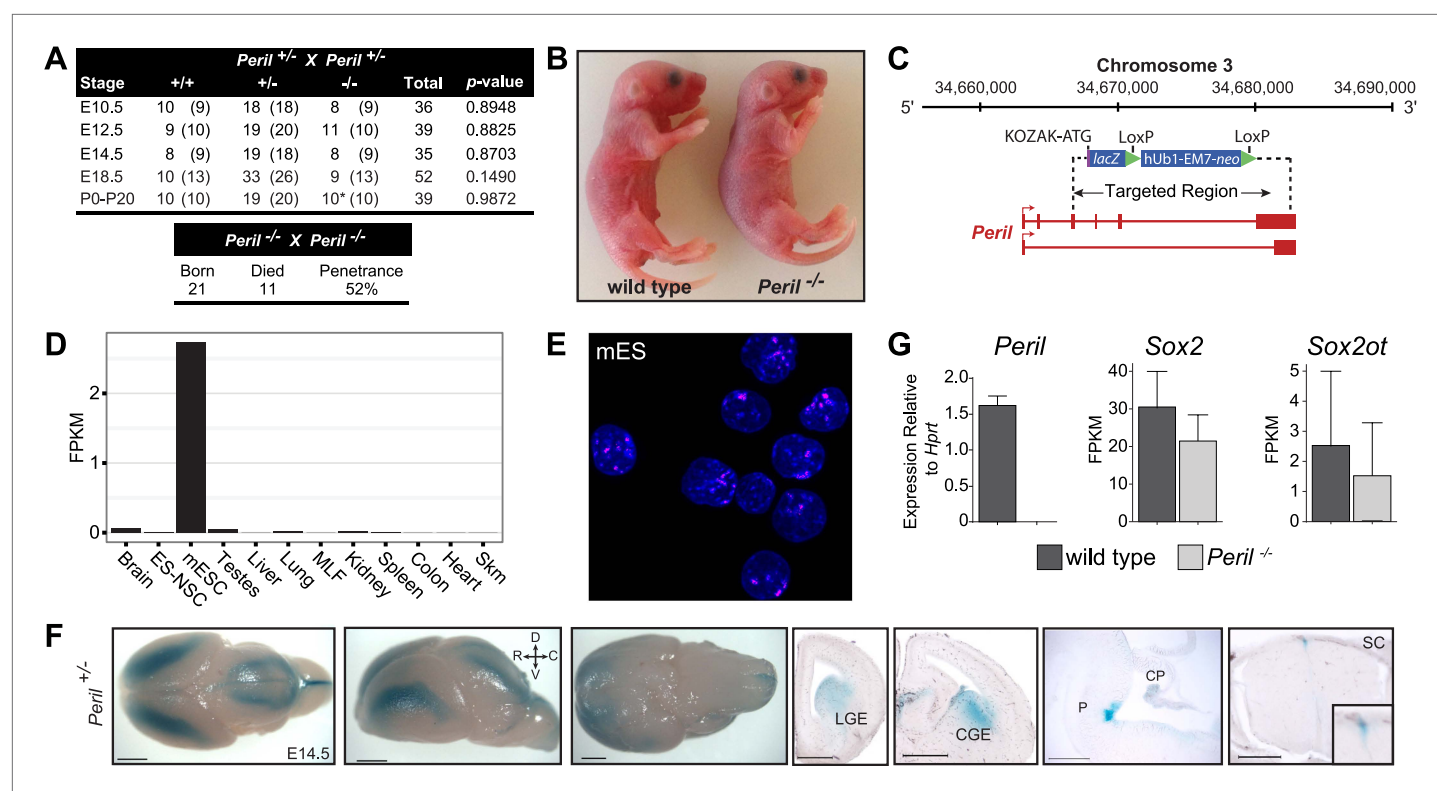
To study the function of the selected 18 lincRNAs in vivo and to resolve, at the cellular level, their expression pattern in different organs, we generated knockout mouse strains for each candidate by replacing the lincRNA gene with a *lacZ* reporter cassette ('Materials and methods', Figure 1—figure supplement 2 for details) (Valenzuela et al., 2003). These new strains, which are the first lincRNA knockout models to incorporate a reporter, more than double the number of models available for investigation and constitute an important resource that will be used to better understand the functional contribution of lincRNAs to mammalian biology.

To assess the requirement for each lincRNA in embryonic development and viability, we examined the progeny from heterozygote intercrosses for all 18 strains. Genotyping of weanlings (21 days old) revealed normal Mendelian segregation of mutant alleles in 15 of the 18 strains (Figure 1D). For the three remaining strains *Peril*, *Mdgt* and *Fendrr*, the progeny of heterozygote intercrosses contained much lower numbers of homozygote mutants than expected. Only 13 *Peril*<sup>-/-</sup> mice (of an expected 32), and 6 *Mdgt*<sup>-/-</sup> mice (of an expected 17) were found at weaning age, indicating that deletion of *Peril* and *Mdgt* leads to reduced viability with >50% and 65% penetrance, respectively (Figure 1D). Closer examination of *Mdgt* pups revealed that homozygous mutants died within 2 weeks after birth. For *Fendrr*, no homozygous mutants were found at weaning age (Figure 1D), indicating that the lethal phenotype for this strain is fully penetrant. Thus, 3 out of the 18 (17%) lincRNA knockout strains generated have a lethal phenotype, confirming that specific lincRNA genes are required for viability.

### *Peril*<sup>-/-</sup> mice have reduced viability and die shortly after birth

To determine the onset of *Peril*<sup>-/-</sup> mice lethality, we monitored survival at both early and late stages of embryonic development. Since normal Mendelian ratios were observed up to E18.5 (Figure 2A), and pups





**Figure 2.** Deletion of *Peril* leads to reduced viability. **(A)** Genotyping results from heterozygote intercrosses (Upper panel) at different developmental stages (\*pups dying within days after birth) and homozygote intercrosses (Lower panel) at birth. The p value is based on  $\chi^2$  test. **(B)** Newborn (P0) *Peril*<sup>-/-</sup> mutants and wild-type littermates. **(C)** *Peril* genomic locus and targeting scheme. **(D)** RNA-Seq expression profile for *Peril* across a panel of mouse tissues and cell types. **(E)** Single-molecule FISH targeting *Peril* in wild type mouse embryonic stem cells (mES). Nuclei were stained with DAPI. **(F)** Whole mount and coronal section *lacZ* stainings reporting expression of *Peril* in the brain and spinal cord of a heterozygote E14.5 embryo. LGE/CGE, Lateral and Caudal Ganglionic Eminence; P, Pons; CP, Choroid Plexus; SC, Spinal Cord; D, dorsal; V, ventral; R, rostral; C, caudal. Scale bars = 1 mm, whole brains; 500  $\mu$ m, sections. **(G)** Relative expression levels of *Peril* as assessed by RT-PCR and expression estimates (FPKM) of the neighboring genes *Sox2* and *Sox2ot* as assessed by RNA-Seq in E18.5 brain of homozygote mutant and wild-type littermates (n = 2 each).

DOI: [10.7554/eLife.01749.006](https://doi.org/10.7554/eLife.01749.006)

The following figure supplements are available for figure 2:

**Figure supplement 1.** *Peril* E18.5 brain differential RNA-Seq analysis.

DOI: [10.7554/eLife.01749.007](https://doi.org/10.7554/eLife.01749.007)

appeared macroscopically normal at birth (**Figure 2B**), we monitored newborns to see if lethality occurred perinatally. We observed that 50% (5/10) of *Peril*<sup>-/-</sup> pups died within 2–20 days after birth (**Figure 2A**). A similar percentage (52%, 11/21) of newborn deaths occurred in the progeny of intercrosses from surviving *Peril*<sup>-/-</sup> mice (**Figure 2A**, lower panel). These results confirmed the reduced viability of *Peril* mutants (~50% penetrance) with pups dying early after birth.

*Peril* is a seven exon transcript derived from a 18.2 Kb genomic locus located ~110 Kb downstream of the pluripotency factor *Sox2* (**Figure 2C**). Using mouse ES cell (mES) cDNA, we were able to clone two distinct isoforms (1,831 bp and 631 bp) (**Figure 2C**). RNAseq expression profiling from a panel of mouse tissues and cell lines showed that *Peril* was highly enriched in mouse ES cells, but also expressed at lower levels in adult brain and testes (**Figure 2D**). The *Peril* transcript is predominantly nuclear as revealed by RNA fluorescence in situ hybridization (FISH) in mouse ES cells (**Figure 2E**). Using the knocked-in *lacZ* reporter in heterozygote embryos, tight temporal and spatial regulation of *Peril* was found in the brain and spinal cord of E14.5 and E18.5 embryos (**Figure 2F**, **Figure 2—figure supplement 1A**).

To identify putative pathways affected by deletion of *Peril*, we performed RNA-sequencing on E18.5 brains from *Peril*<sup>-/-</sup> and wild-type littermates (n = 2). Analysis of *Peril* expression levels confirmed the deletion of *Peril* (**Figure 2G**). Expression of the pluripotency factor *Sox2* and its overlapping

noncoding RNA *Sox2ot* were not significantly affected in the knockout brains (**Figure 2G**). Differentially regulated genes and gene set enrichment analysis (GSEA) revealed downregulation of genes involved in cell cycle regulation, energy metabolism, immune response, and mRNA and protein processing in *Peril*<sup>-/-</sup> brains relative to wild-type controls (**Figure 2—figure supplement 1B and C**). This indicates that the transcriptional programs within the brain are in fact affected by deletion of *Peril*. Yet, it remains unclear if these changes underlie the observed lethality.

### ***Fendrr*<sup>-/-</sup> is perinatal lethal due to defects in multiple organs**

*Fendrr* is a 2,380 bp transcript consisting of six exons. It is transcribed from a bidirectional promoter shared with the protein coding gene *Foxf1a*, located 1,354 bp from its transcriptional start site (**Figure 3A**). An orthologous human *FENDRR* (*LINC-FOXF1*) transcript expressed from a syntenic region was identified within our catalog of human lincRNAs (*Cabili et al., 2011*). We previously demonstrated that this lincRNA is predominantly nuclear and physically associates with the PRC2 Polycomb complex (*Khalil et al., 2009*).

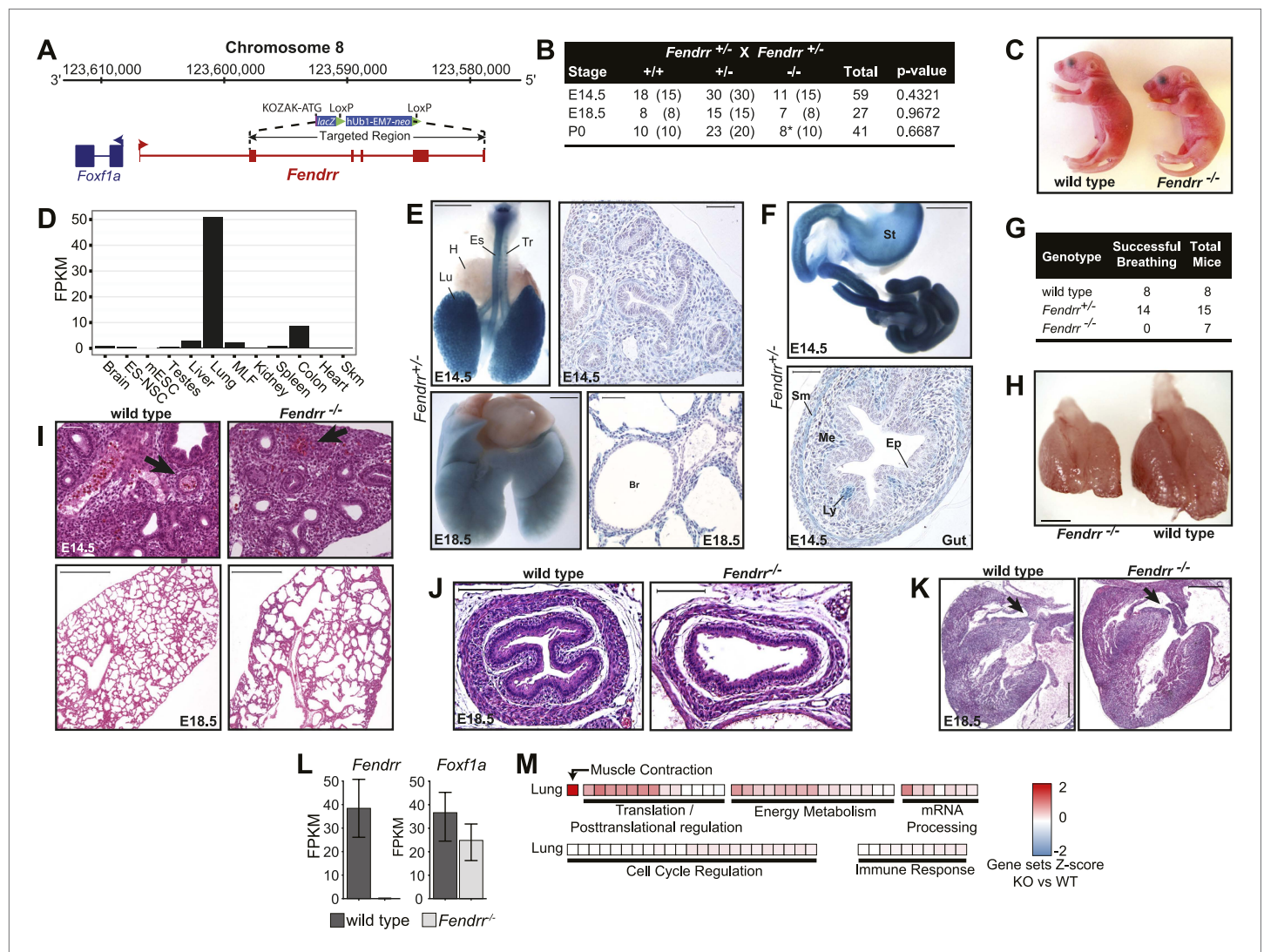
To determine the onset of *Fendrr* lethality, we monitored the survival of embryos at both early and late stages of embryonic development, as well as that of newborn pups. Normal Mendelian ratios were found at both E14.5 and E18.5 (**Figure 3B**), with embryos appearing macroscopically normal prior to birth (**Figure 3C**), suggesting that the lethality most likely occurred after birth. Importantly, we observed 8 *Fendrr*<sup>-/-</sup> mutant newborns (P0) (**Figure 3B**), all of which died within 24 hr, indicating perinatal lethality.

During the course of our study, another group generated a *Fendrr* loss-of-function mouse model by replacing the first exon with multiple transcriptional termination sequences (*Grote et al., 2013*). In contrast to the results presented here, Grote et al. observed lethality at E13.75 due to heart and body wall (omphalocele) defects. When analyzing E14.5 embryos, we found no resorbed embryos or omphalocele in our *Fendrr* homozygous mutants (**Figure 3—figure supplement 1**). Although, both studies used similar genetic background strains, a possible explanation for this discrepancy may be found in the distinct targeting strategy used to remove the *Fendrr* gene. Regardless, both studies confirm that loss of *Fendrr* is lethal in mice.

Grote et al. observed *Fendrr* expression to be restricted to nascent lateral plate mesoderm and did not detect it in any other tissue. Using RNA-Seq expression profiling from adult mouse tissues and cell lines, we however found that *Fendrr* is expressed at high levels in the adult lung, and lower levels are detectable in colon, liver, spleen and brain (**Figure 3D**). Analysis of the knocked-in *lacZ* reporter in E14.5 and E18.5 embryos confirmed expression of *Fendrr* in these tissues as well as in the trachea and all along the gastrointestinal tract (**Figure 3E,F**). Interestingly, in the developing respiratory and digestive tracts at E14.5 and E18.5, expression of *Fendrr* is restricted to the pulmonary mesenchyme surrounding the bronchiolar epithelial cells (**Figure 3E**, right panels); in mesenchymal cells of the developing mucosa; the muscularis externa of the gut and in lymphoid aggregations within the gut's mucosa (**Figure 3F**, bottom panel), all of which are derived from the lateral plate mesoderm.

Perinatal lethality is often associated with respiratory failure. Since the highest expression levels of *Fendrr* are found in the lungs, we evaluated initiation of breathing in surgically delivered E18.5 embryos. After cleaning of their airways, all *Fendrr*<sup>-/-</sup> embryos analyzed either failed to breathe or gasped and stopped breathing within 5 hr ( $n = 7$  *Fendrr*<sup>-/-</sup>, **Figure 3G**). In contrast, respiration initiated normally and was maintained for all but one of the heterozygote and wild-type embryos ( $n = 15$  *Fendrr*<sup>+/-</sup> and  $n = 8$  wild type). *Fendrr*<sup>-/-</sup> lungs at the pseudoglandular stage (E14.5) were hypoplastic compared to wild type (**Figure 3H**), and histological evaluation of the lungs revealed a decrease in the number and organization of pulmonary arteries, and a general failure of vasculogenesis within the lungs of the *Fendrr*<sup>-/-</sup> mutants compared to wild type ( $n = 3$  *Fendrr*<sup>-/-</sup> and  $n = 3$  wild type; **Figure 3I**, upper panels). At E18.5, *Fendrr*<sup>-/-</sup> lungs appeared to have fewer but larger alveoli ( $n = 3$  *Fendrr*<sup>-/-</sup> and  $n = 3$  wild type; **Figure 3I**, lower panels). Together, these results suggest that respiratory failure observed at birth in *Fendrr*<sup>-/-</sup> mice could be due to a lung maturation and vascularization defect.

We also observed expression of *Fendrr* in the esophagus and gut. We observed thinning of the mesenchymal layer of the mucosa and external smooth muscle layers in the esophagus at E18.5 ( $n = 3$  *Fendrr*<sup>-/-</sup> and  $n = 3$  wild type; **Figure 3J**). Although we did not observe *Fendrr* expression in the heart at E14.5, E18.5, or postnatally, we did observe intraventricular septal heart defects prior to birth (E18.5) ( $n = 3$  *Fendrr*<sup>-/-</sup> and  $n = 3$  wild type; **Figure 3K**). In accordance with *Fendrr* having a previously described role in lateral plate mesoderm (*Grote et al., 2013*), our results suggest a more general role



**Figure 3.** *Fendrr*<sup>-/-</sup> pups have multiple defects in lung, heart and gastrointestinal tract. **(A)** *Fendrr* locus and targeting strategy. Arrows indicate location of the primers used for genotyping. **(B)** Genotyping results from heterozygote intercrosses at embryonic stages E14.5, E18.5 and at birth (P0). The p value is based on  $\chi^2$  test. (\*) All newborns died within 24 hr after birth. **(C)** *Fendrr*<sup>-/-</sup> E18.5 embryos and wild-type littermates. **(D)** RNA-Seq expression profile for *Fendrr* across a panel of mouse tissues and cell types. **(E and F)** *lacZ* reporter stained organs and sections showing expression of *Fendrr* in specific regions of the lung (Lu), trachea (Tr) and esophagus (Es), but not in heart (H) in E14.5 and E18.5 embryo **(E)** and in the gut and stomach (St) **(F)**. Sm, smooth muscle; Ep, Epithelia; Me, Mesenchyme; Ly, Lymphoid aggregates. Scale bars = 1 mm whole organ, 200  $\mu$ m sections. **(G)** Number of E18.5 embryos successfully breathing after surgical delivery. **(H)** Size difference of *Fendrr*<sup>-/-</sup> lungs at E14.5 compared to wild-type littermates ( $n = 3$  each). **(I–K)** Representative hematoxylin and eosin (H&E) stained sections showing unstructured vessels (arrow) in E14.5 *Fendrr* mutant lungs compared to wild type littermates ( $n = 3$ ) **(I)**, upper panels), alveolar defects at E18.5 **(I)**, lower panel), thinner mesenchymal layer of the mucosa and external smooth muscle layer of the esophagus **(J)** and ventricular septal defects in the heart **(K)** of *Fendrr*<sup>-/-</sup> E18.5 embryos compared to wild type ( $n = 3$ ). Scale bars = 500  $\mu$ m, 100  $\mu$ m for esophagus. **(L)** RNA-Seq expression levels of *Fendrr* and the neighboring coding gene *Foxf1a* in E18.5 lung of homozygote mutant and wild-type littermates ( $n = 2$  each). **(M)** GSEA of *Fendrr*<sup>-/-</sup> vs wild-type E18.5 lung and total brain RNA. Each tile is a significant ( $q < 0.001$ ; Mann-Whitney, BH) gene set from the Reactome collection at mSigDB, based on the *Fendrr*<sup>-/-</sup>/wild-type ranking of test-statistic values from a Cuffdiff2 differential analysis. Tiles are shaded based on the z-score of the test-statistic for genes within the given gene set, relative to all genes for a given condition to show direction of expression relative to wild-type.

DOI: 10.7554/eLife.01749.008

The following figure supplements are available for figure 3:

**Figure supplement 1.** *Fendrr*<sup>-/-</sup> embryos don't have an omphalocele.

DOI: 10.7554/eLife.01749.009

**Figure supplement 2.** *Fendrr*, *linc-Brn1b*, and *Peril* do not act as cis-enhancer elements.

DOI: 10.7554/eLife.01749.010



for *Fendrr* in regulating the proper differentiation of mesenchyme-derived tissues across several organ systems.

We next investigated if neighboring gene expression is perturbed by the deletion of *Fendrr*. We harvested lungs from E14.5 *Fendrr*<sup>-/-</sup> embryos and wild-type littermates (*n* = 2) and performed differential RNA-Seq analyses. A loss of *Fendrr* expression in knockout relative to wild-type lungs confirmed deletion of *Fendrr* (**Figure 3L**). No significant change in the expression of the *Foxf1a* protein coding gene was observed in the *Fendrr*<sup>-/-</sup> lung. Furthermore, genes within  $\pm 1$  Mb of the *Fendrr* locus were not significantly differentially expressed any more than background levels of local enrichment (**Figure 3—figure supplement 2A**, *p* < 0.087; bootstrapped from random 2 Mb genomic intervals), suggesting that the *Fendrr* gene does not act as a local enhancer. GSEA identified gene sets involved in muscle differentiation and contraction as the most significant sets misregulated in *Fendrr*<sup>-/-</sup> lungs compared to wild type (**Figure 3M**). This agrees with our identification of defects in the lung vasculature of the *Fendrr*<sup>-/-</sup> mice. Further studies will be needed to understand how specific changes in gene-expression patterns upon deletion of *Fendrr* contribute to the observed defects and perinatal lethality.

## Ablation of distinct lincRNA genes affects mouse growth

To determine if deletion of our candidate lincRNAs affects normal development and growth postnatally, each strain was examined for gross morphological abnormalities and body weight (BW) measurements were taken. Homozygote mutant mice with heterozygote and wild-type littermate controls were compared over a 7–10 week period. *Mdgt*<sup>-/-</sup> pups displayed a severe growth retardation phenotype, which may contribute to their lethality (**Figure 4C,D**). A week after birth, *Mdgt*<sup>-/-</sup> pups were already significantly smaller than heterozygote and wild-type littermates (**Figure 4C**), with females 60% smaller (*n* = 5, *p* < 0.00001) than wild-type littermates (*n* = 10) and males 32% smaller (*n* = 3, *p* < 0.05), suggesting a sex bias in this phenotype. In *Mdgt*<sup>-/-</sup> survivors, this retarded growth persisted up to 10 weeks in females, which were still 37% smaller than wild types (**Figure 4D**, *p* < 0.0006). The defect, although milder, also persisted in male mutants up to at least 8 weeks.

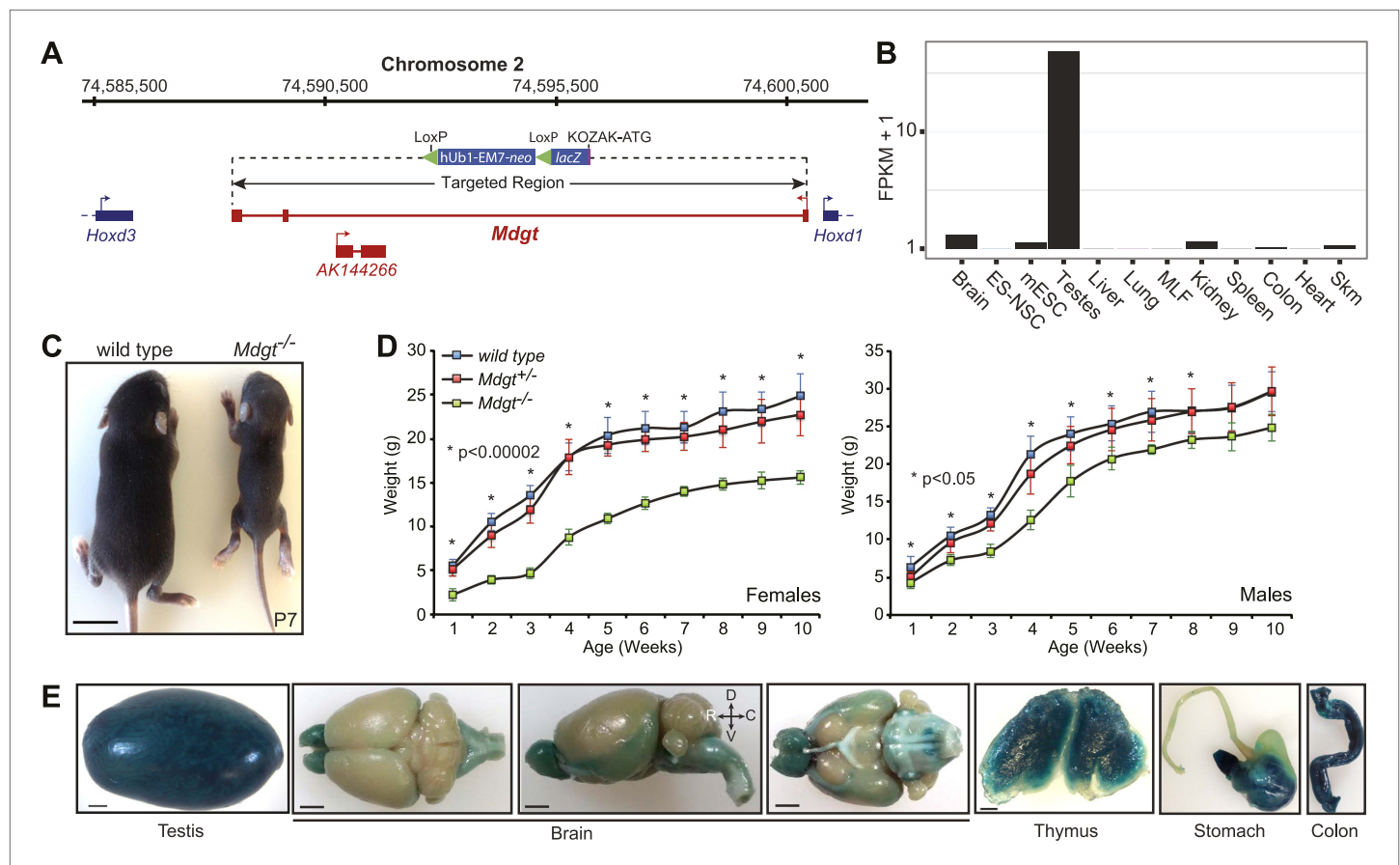
*Mdgt* is a 443 bp transcript consisting of three exons transcribed from a bidirectional promoter shared with the homeobox gene *Hoxd1*. The first exon for *Mdgt* is located only 84 bp away from the *Hoxd1* transcription start site. Another noncoding transcript, AK144266, is contained within an intron of *Mdgt* and appears to be expressed from the opposite strand (**Figure 4A**). In our targeting strategy, the entire *Mdgt* locus was deleted, including the AK144266 transcript (**Figure 4A**). Expression estimates obtained from RNA-Seq revealed that *Mdgt* is highly expressed in testes, and more moderately in brain, kidney, colon and skeletal muscles (**Figure 4B**). *lacZ* staining confirmed expression in those tissues and also revealed expression in stomach and thymus (**Figure 4E**).

In addition to *Mdgt*, stunted growth defects were also observed in three other strains (**Figure 4—figure supplement 1**). 3 weeks after birth, *linc-Pint*<sup>-/-</sup> mice were noticeably smaller than wild-type littermates (**Figure 4—figure supplement 1A**). By 7 weeks of age, females had 22% reduced body weight (*n* = 5, *p* < 0.001) compared to wild type (*n* = 4), whereas males were 29% smaller (*n* = 3 *linc-Pint*<sup>-/-</sup>; *n* = 3 wild type, *p* < 0.05). A decrease in body weight was also observed by three weeks of age in *linc-Brn1b*<sup>-/-</sup> (29%, *p* < 0.013, *n* = 3) and *Peril*<sup>-/-</sup> mice (35%, *p* < 0.0003, *n* = 4) when compared to wild-type littermates (*n* = 3) (**Figure 4—figure supplement 1B,C**). Together, these results indicate that specific lincRNAs can be important for normal body weight and growth.

## Identification of developmentally regulated human lincRNA orthologs

In order to determine whether loss of our candidate lincRNAs could lead to other developmental defects, we examined additional mutant strains that did not show perinatal lethality. Building on prior evidence that lincRNAs are abundantly expressed and spatio-temporally regulated within the brain during development and adulthood (Mercer et al., 2008; Qureshi et al., 2010; Cabili et al., 2011; Ramos et al., 2013), we concentrated on the candidate lincRNAs that were expressed in the brain or in neural progenitors. Twelve of the lincRNAs for which we have established knockout strains exhibited such expression profile (**Figure 1B**). To focus on lincRNA candidates of potential functional relevance in neuronal development, we used syntenic orthology (TransMap [Zhu et al., 2007]) and RNA-sequencing to select those with identifiable human orthologs, and whose expression was regulated during in vitro neural differentiation.

Briefly, transcripts expressed along a time course of ES-derived human neural stem cells (NSC) differentiation (Goff et al., 2009) (**Figure 5A**) were assembled and aggregated with an existing compendium of RNA-Seq data, using our previously described lincRNA discovery pipeline (Cabili et al., 2011).



**Figure 4.** *Mdg1*<sup>-/-</sup> surviving mice have growth defects. **(A)** *Mdg1* genomic locus and targeting scheme. **(B)** RNA-Seq expression profile for *Mdg1* across a panel of mouse tissues and cell types. **(C)** Representative example showing the reduced size of *Mdg1*<sup>-/-</sup> pups compared to wild type 7 days after birth (P7). **(D)** Body weight (g) measurements over a 10 weeks period show growth retardation in both female and male *Mdg1*<sup>-/-</sup> mice compared to wild type and *Mdg1*<sup>+/-</sup> littermates (Females: n = 5 *Mdg1*<sup>-/-</sup>, n = 15 *Mdg1*<sup>+/-</sup> and n = 10 wild types; Males: n = 3 *Mdg1*<sup>-/-</sup>, n = 10 *Mdg1*<sup>+/-</sup> and n = 10 wild types). Significant p values at each time point are indicated (\*). **(E)** Whole mount *lacZ* stainings reporting expression of *Mdg1* in adult tissues of heterozygote mutant mice. Scale bars = 1 mm, testis, thymus; 2 mm, brain, stomach, colon.

DOI: 10.7554/eLife.01749.011

The following figure supplements are available for figure 4:

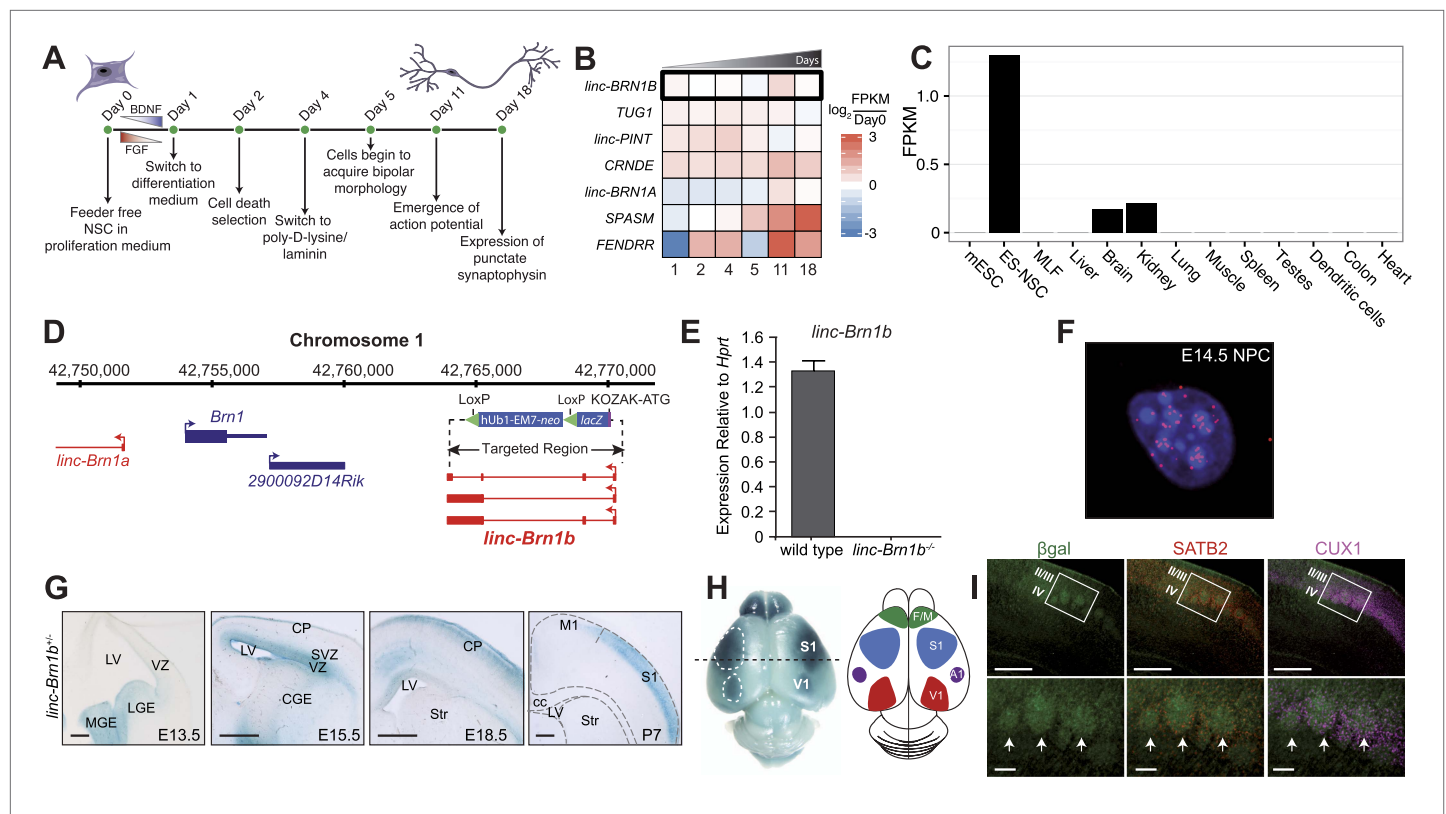
**Figure supplement 1.** Growth retardation phenotype in lincRNA knockout strains.

DOI: 10.7554/eLife.01749.012

The resulting lincRNA catalog contained 24,737 distinct, high-quality transcript reconstructions corresponding to 14,259 human lincRNA genes. We observed 769 lincRNA genes with significant differential expression ( $q < 0.01$ ; Cuffdiff2) between any two adjacent time points during NSC differentiation. With a false discovery rate of 1% (Figure 5—figure supplement 1), 302 of these were significantly induced relative to day 0 (Figure 5—figure supplement 1). This approach revealed that 7 lincRNAs from our mouse knockout strains have human orthologs that are dynamically induced during in vitro human neuronal differentiation (Figure 5B). Interestingly, two of these, *linc-Brn1a* and *linc-Brn1b*, were almost exclusively expressed in NSCs as determined by RNA-Seq (Figure 1B). These lincRNAs reside in the genomic region of *Brn1* (*Pou3f3*), a well-studied transcription factor involved in cortical development (McEvilly, 2002; Sugitani et al., 2002; Dominguez et al., 2012). We thus focused on this locus, starting with *linc-Brn1b*, as it is not transcribed from a bidirectional promoter and therefore deletion does not disrupt the *Brn1* promoter.

### *linc-Brn1b* expression is spatio-temporally restricted in the developing cerebral cortex

Using RNA-Seq data, we observed *linc-Brn1b* to be predominantly expressed in ES-derived NSCs with additional moderate expression in the adult brain and kidney (Figure 5C). The *linc-Brn1b* gene is a



**Figure 5.** *linc-Brn1b* is spatio-temporally regulated during corticogenesis. **(A)** Schematic overview of the in vitro human neural stem cell differentiation protocol. RNA was collected at the indicated time points and sequenced to identify significantly differentially expressed lincRNA human orthologs. **(B)** Heatmap of  $\log_2$  ratios to undifferentiated (Day 0) human neural stem cells for 7 of the 20 lincRNAs selected for deletion with a significant ( $q < 0.01$ ; Cuffdiff2) increase in expression during differentiation. **(C)** RNA-Seq expression profile for *linc-Brn1b* across a panel of mouse tissues and cell types. **(D)** *linc-Brn1b* genomic locus and targeting strategy. **(E)** qRT-PCR confirmation of the genotype for both heterozygotes (+/-) and homozygous null (-/-) mutants. **(F)** Single-molecule RNA FISH targeting *linc-Brn1b* in wild-type E14.5 cortical neurospheres. **(G)** *lacZ* staining shows expression profile of *linc-Brn1b* at different embryonic (E13.5, E15.5 and E18.5) and early postnatal stages (P7) in *linc-Brn1b*<sup>+/-</sup> telencephalon. *lacZ* expression in neural progenitors of both ventral telencephalon (lateral ganglionic eminence, LGE and medial ganglionic eminence, MGE) (E13.5) and dorsal telencephalon (ventricular zone, VZ and subventricular zone, SVZ) (E15.5) is detected. Restricted expression in the upper cortical layers is observed in both E18.5 and P7 cortex. **(H)** Whole mount *lacZ* staining of P7 *linc-Brn1b*<sup>+/-</sup> brain shows distinct *linc-Brn1b* expression in primary somatosensory (S1) and visual (V1) cortical areas. **(I)**  $\beta$ -galactosidase immunofluorescence on coronal sections of P7 *linc-Brn1b*<sup>-/-</sup> cortex shows *linc-Brn1b* expression in layer IV of the somatosensory area (white boxes), specifically within the barrel structures (white arrows), as determined by co-staining with the upper layer markers SATB2 and CUX1. Abbreviations: LV, lateral ventricle; LGE, lateral ganglionic eminence; MGE, medial ganglionic eminence; CGE, caudal ganglionic eminence; CP, cortical plate; SVZ, subventricular zone; VZ, ventricular zone; Str, striatum; M1, primary motor area; S1, primary somatosensory area; cc, corpus callosum; V1, primary visual cortex; A1, primary auditory cortex; F/M, frontal motor cortex. Scale bars: 500  $\mu$ m (**G**), (I, upper panels); 100  $\mu$ m (I, lower panels).

DOI: [10.7554/eLife.01749.013](https://doi.org/10.7554/eLife.01749.013)

The following figure supplements are available for figure 5:

**Figure supplement 1.** Differential expression of protein coding genes, lncRNAs and neuronal markers during human ES-derived neuronal differentiation time course.

DOI: [10.7554/eLife.01749.014](https://doi.org/10.7554/eLife.01749.014)

~3 Kb transcript derived from a 6.8 Kb genomic locus approximately 10 Kb downstream of the *Brn1* protein coding gene. To generate the *linc-Brn1b* knockout mice, we targeted the entire *linc-Brn1b* locus (**Figure 5D**). Complete ablation was confirmed by qRT-PCR (**Figure 5E**) using adult brain cDNA as template. RNA-FISH in mouse E14.5 neural progenitor cells (NPCs) isolated from the cerebral cortex demonstrated that the *linc-Brn1b* transcript is predominantly nuclear with moderate cytoplasmic expression (**Figure 5F**).

While identified as a lincRNA expressed in adult brain, the spatio-temporal distribution of *linc-Brn1b* during brain development is not known. Therefore, we used *lacZ* expression from the *linc-Brn1b* locus in heterozygote mutants to define its expression in vivo. We found that *linc-Brn1b* was expressed within neural progenitors of both the ventral and dorsal telencephalon, as early as E13.5 (**Figure 5G**).

Detailed characterization of expression in the germinal zones of the dorsal telencephalon showed that by E15.5, it was strongly expressed in progenitors of both the ventricular zone (VZ) and the subventricular zone (SVZ) of the developing cerebral cortex, and by E18.5 showed restricted expression in the developing upper cortical layers. In addition, whole-mount *lacZ* staining at P7, showed a specific areal distribution for *linc-Brn1b* within both the primary somatosensory cortex and the primary visual cortex (**Figure 5H**). Within the somatosensory cortex, *linc-Brn1b* was expressed in the barrel structures of the posteromedial barrel subfield (PMBSF) (**Figure 5I**); a highly organized region of cortical projection neurons that receives afferent connections from the thalamus and is responsible for the coordination of sensory inputs from the rodent vibrissae (**Petersen, 2007**).

Collectively, the timing and location of *linc-Brn1b* expression within the developing cortex suggests a potential role for *linc-Brn1b* in area-specific development of distinct classes of projection neurons.

### Loss of *linc-Brn1b* results in specific reduction in the number of intermediate progenitor cells (IPCs) in the cerebral cortex

To investigate the consequences of genetic deletion of *linc-Brn1b* during development of the telencephalon, we performed RNA-Seq on E13.5 and E15.5 whole telencephalons and P7 whole brains ( $n = 2$ ) of wild type and *linc-Brn1b*<sup>-/-</sup> mice. At all time points, RNA-Seq analyses identified a statistically significant reduction (~50%) in expression of the neighboring *Brn1* protein coding gene in *linc-Brn1b*<sup>-/-</sup> brains relative to wild type (E15.5 shown in **Figure 6A**, **Figure 6—figure supplement 1A,B**). We observed a similar decrease in BRN1 protein expression between wild type and *linc-Brn1b*<sup>-/-</sup> E14.5 cortical-derived neurospheres and E15.5 whole cortex samples (**Figure 6B**). In contrast, upon knockout of *linc-Brn1b*, we observed a significant increase in expression of *linc-Brn1a* ( $p < 0.01$ ; Cuffdiff2), which shares a bidirectional promoter with *Brn1* (**Figure 6A**). This suggests opposing regulatory effects on the neighboring lincRNA and protein coding genes upon deletion of *linc-Brn1b*.

Gene set enrichment analysis (GSEA) revealed a significant decrease in genes associated with cellular growth and proliferation in *linc-Brn1b*<sup>-/-</sup> relative to wild type at all three time points sampled and an increase in gene sets positively correlated with neuronal maturation (**Figure 6C**). These results are consistent with a predicted role for *linc-Brn1b* derived from our initial GBA analyses (**Figure 1C**) and suggest a function for *linc-Brn1b* in the differentiation of neural progenitors and neurogenesis in the telencephalon.

To investigate potential abnormalities of *linc-Brn1b*<sup>-/-</sup> mutants, we concentrated on the dorsal telencephalon (the developing cerebral cortex). First, to quantify the proliferation of cortical progenitors, we evaluated the expression of the mitotic marker phosphorylated histone H3 (pH3) in E15.5 wild type and *linc-Brn1b*<sup>-/-</sup> cortices (**Figure 6D**). Quantification of pH3<sup>+</sup> cells revealed a significant and specific reduction in the relative percentage of pH3<sup>+</sup> intermediate progenitor cells within the SVZ (36.88% decrease;  $p < 3.93 \times 10^{-5}$ ; Student's *t* test) in *linc-Brn1b* mutant vs control cortices (**Figure 6E**). The observed abnormalities were specific to intermediate progenitors, as there was no statistically significant change in the proliferation rate of apical progenitors within the VZ (**Figure 6E**).

In agreement with a reduction in intermediate progenitors of the cortical SVZ we found decreased expression of the markers TBR2 (**Figure 6F**), Cux2 (**Figure 6G**), and Svet1 (**Figure 6H**) (**Tarabykin et al., 2001; Zimmer, 2004; Molyneaux et al., 2007; Arnold et al., 2008; Cubelos et al., 2008; Sessa et al., 2008**) in *linc-Brn1b*<sup>-/-</sup> mutants.

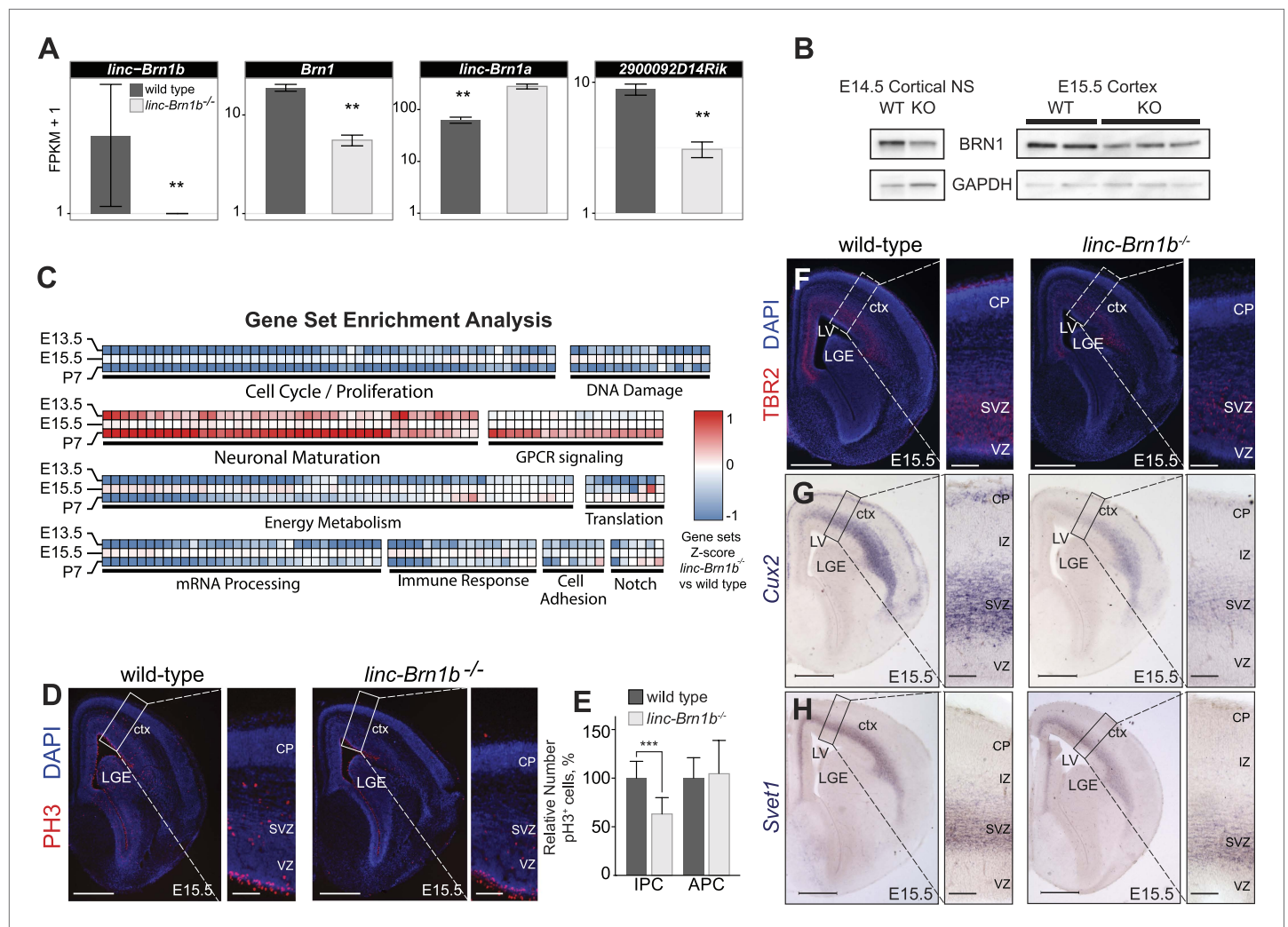
Together, the data indicate that *linc-Brn1b* knockout mice exhibit a decrease in proliferation of cortical progenitors in vivo, and that this defect is restricted to a specific subpopulation of progenitors in the SVZ of the developing cortex.

### Cortical lamination is abnormal in *linc-Brn1b*<sup>-/-</sup> mice

To investigate whether the observed decrease in SVZ progenitors resulted in an overall reduction in cortical thickness, we measured *linc-Brn1b*<sup>-/-</sup> mutants and wild-type littermates at P7. The distance between the pia and the white matter was measured at matched medio-lateral and rostro-caudal locations ( $n = 17$  wild-type sections and  $n = 10$  *linc-Brn1b*<sup>-/-</sup> sections). We reproducibly observed a significant 6.24% decrease in the total thickness of the neocortex in the mutants, relative to wild type ( $1317.31 \mu\text{m} \pm 20.03$  in *linc-Brn1b*<sup>-/-</sup> vs  $1404.90 \mu\text{m} \pm 15.07$  in wild type;  $p < 0.002$ , Student's *t* test) (data not shown).

To define whether generation of projection neuron subtypes was affected in the absence of *linc-Brn1b*, cortices from *linc-Brn1b*<sup>-/-</sup> and wild-type littermates were collected at P7 and immunostained for markers of different projection neuron classes: CUX1, a marker of upper layer II/III and IV callosal





**Figure 6.** *linc-Brn1b*<sup>-/-</sup> mice demonstrate defects in proliferation of IPCs. (A) RNA-Seq expression estimates from E15.5 wild type and *linc-Brn1b*<sup>-/-</sup> telencephalon for *linc-Brn1b*, the protein coding gene *Brn1*, and neighboring genes *linc-Brn1a* and *2900092D14Rik*. (B) Western blots of wild type and *linc-Brn1b*<sup>-/-</sup> E14.5 cortical neurospheres (NS), and E15.5 cortices. (C) GSEA of *linc-Brn1b*<sup>-/-</sup> vs wild type in E13.5 and E15.5 telencephalon and P7 whole brain (as described in Figure 3). (D and E) Immunofluorescence staining for the mitotic marker phosphorylated histone H3 (pH3) in E15.5 coronal sections of cortex (D) and pH3<sup>+</sup> cell counts of apical progenitors (APC) and intermediate progenitors (IPC) (E). (F-H) Immunofluorescence staining for TBR2 (F), and in situ hybridization for *Cux2* (G) and *Svet1* (H) in wild type and *linc-Brn1b*<sup>-/-</sup> E15.5 cortex show that mutant mice have decreased expression of SVZ intermediate progenitor markers. Scale bars = 500 μm (D-H). \*\*p < 0.01, \*\*\*p < 0.001; Student's t test. LGE, lateral ganglionic eminence; LV, lateral ventricle; ctx, cortex; CP, cortical plate; VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone.

DOI: 10.7554/eLife.01749.015

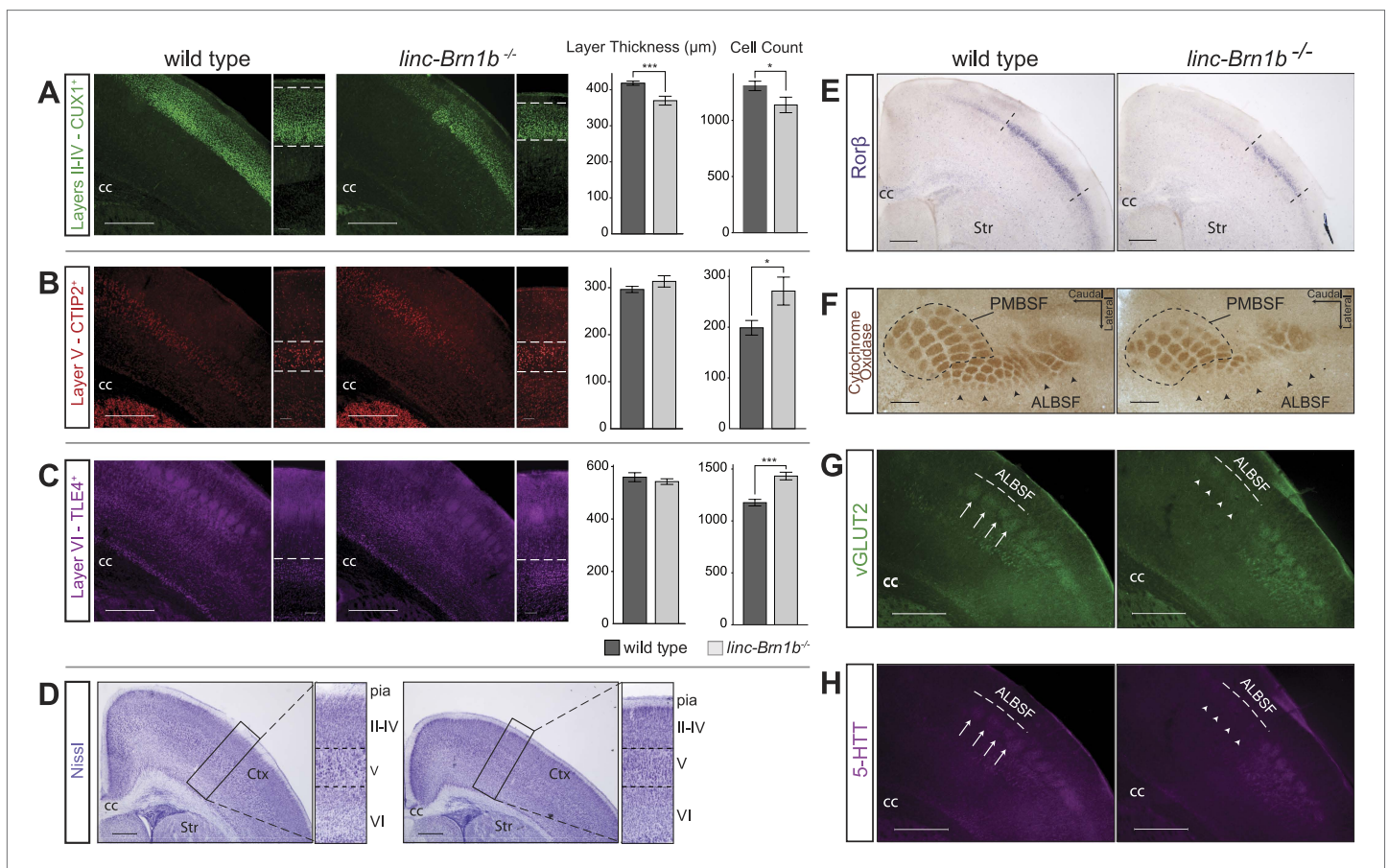
The following figure supplements are available for figure 6:

**Figure supplement 1.** *Brn1* transcript is reduced in *linc-Brn1b*<sup>-/-</sup> E13.5 telencephalon and P7 brain.

DOI: 10.7554/eLife.01749.016

projection neurons (Zimmer, 2004) (Figure 7A), CTIP2, a marker of layer V subcerebral projection neurons (Arlotta et al., 2005) (Figure 7B), and TLE4, a marker of layer VI corticothalamic projection neurons (Koop et al., 1996; Chen, 2005; Molyneux et al., 2005) (Figure 7C). Measurements of the thickness of each cortical layer showed a distinct reduction in the upper layer II/III–IV (11.62% reduction in absolute thickness in *linc-Brn1b*<sup>-/-</sup> vs wild type; Figure 7A, p < 0.0004; Student's t test), which was also detectable by histological Nissl staining (Figure 7D). In contrast, no significant change in the absolute thickness of either layer V (Figure 7B, p < 0.18; Student's t test) or VI (Figure 7C, p < 0.54; Student's t test) was detected.

In order to determine whether the number of specific classes of neurons within layers was affected in the mutant cortex, we quantified the percentage of CUX1<sup>+</sup>, CTIP2<sup>+</sup> and TLE4<sup>+</sup> neurons. In agreement with a reduction in the thickness of layer II/III–IV, we observed that the number of CUX1<sup>+</sup> callosal



**Figure 7.** Abnormal cortical lamination and disruption of the barrel cortex in *linc-Brn1b*<sup>-/-</sup> mice. **(A)** Immunofluorescence staining and quantification on coronal sections of P7 wild type and *linc-Brn1b*<sup>-/-</sup> cortices for upper layer II–IV marker CUX1 show a significant reduction in absolute layer thickness and total number of CUX1<sup>+</sup> projection neurons in *linc-Brn1b*<sup>-/-</sup> mice. **(B and C)** Immunofluorescence staining and quantification on coronal sections of P7 wild type and *linc-Brn1b*<sup>-/-</sup> cortices for layer V marker CTIP2 **(B)**, and for layer VI marker TLE4 **(C)** show not significant change in the total thickness of layer V and VI, but significant increase in the number of CTIP2<sup>+</sup> subcerebral projection neurons and TLE4<sup>+</sup> corticothalamic neurons in *linc-Brn1b*<sup>-/-</sup> mice. **(D)** Nissl staining of coronal sections of P8 wild type and *linc-Brn1b*<sup>-/-</sup> cortices shows overall cortical reduction and specific decrease in upper layer thickness in *linc-Brn1b*<sup>-/-</sup> mice. **(E)** *Rorb* in situ hybridization on coronal sections of P7 wild type and *linc-Brn1b*<sup>-/-</sup> primary somatosensory cortex. **(F)** Cytochrome oxidase c activity on sections spanning barrel cortex shows reduction in the anteriolateral barrel subfield (ALBSF) in *linc-Brn1b*<sup>-/-</sup> mice. The loss of barrels in the ALBSF and their general disorganization are confirmed by immunofluorescence staining for vGLUT2 **(G)** and 5-HTT **(H)**. The full arrows indicate individual barrels and arrowheads point to corresponding barrels that are absent in the *linc-Brn1b*<sup>-/-</sup> brains. \**p*<0.05, \*\**p*<0.01, \*\*\**p*<0.001; Student's *t* test. Scale bars = 500 μm **(A–E)** and **(G–H)**, 100 μm **(F)**. ctx, cortex; cc, corpus callosum; str, striatum; PMBSF and ALBSF, posteromedial and anteriolateral barrel subfield.

DOI: 10.7554/eLife.01749.017

projection neurons was decreased by 12.97% in the mutant compared to wild type (**Figure 7A**, mean = 1137 ± 69 CUX1<sup>+</sup> cells in *linc-Brn1b*<sup>-/-</sup> vs mean = 1307 ± 52 in wild type; *p*<0.03; Student's *t* test). Conversely, despite the unchanged thickness of layers V and VI, the number of CTIP2<sup>+</sup> subcerebral projection neurons (**Figure 7B**, mean = 271 ± 28 in *linc-Brn1b*<sup>-/-</sup> vs mean = 199 ± 14 in wild type; *p*<0.016; Student's *t* test) and TLE4<sup>+</sup> corticothalamic projection neurons (**Figure 7C**, mean = 1432 ± 37 in *linc-Brn1b*<sup>-/-</sup> vs mean = 1176 ± 31 in wild type; *p*<5.9 × 10<sup>-5</sup>; Student's *t* test) neurons were increased by 36.39% and 21.77% in the *linc-Brn1b*<sup>-/-</sup> mice, respectively. These data suggest the possibility that a subset of upper layer progenitors abnormally generate deep layer neurons in place of CUX1<sup>+</sup> upper layer neurons in the *linc-Brn1b*<sup>-/-</sup> mutant.

The selective reduction in the generation of upper layer cortical neurons is interesting, considering that these neurons are principally derived from the expansion of intermediate progenitors, which are reduced in *linc-Brn1b*<sup>-/-</sup> mice. Taken together, these results indicate that *linc-Brn1b* is required for proper generation of cortical neurons, in particular projection neurons of layer II/III–IV.

## Disorganization of the barrel cortex in *linc-Brn1b*<sup>-/-</sup> mice

Whole mount analysis of *linc-Brn1b* areal distribution showed high levels of expression in primary somatosensory cortex, an area that in rodents receives thalamocortical afferents relaying sensory information from the mustacial vibrissae (Petersen, 2007). Given the expression of *linc-Brn1b* within this region, and the requirement for proper specification of upper layer neurons, we investigated whether *linc-Brn1b* is required for the proper development of the somatosensory cortex and organization of the barrel structures. In situ hybridization against the barrel cortex marker RAR-related orphlin beta (*Rorb*) (Jabaudon et al., 2012), on matched coronal sections of P7 wild type and knockout mouse cortices demonstrated a reduction in the total size of the barrel cortex in the *linc-Brn1b*<sup>-/-</sup> mice (Figure 7E) with a more pronounced loss of *Rorb*<sup>+</sup> neurons at the medial edge.

Histochemical staining of cytochrome oxidase C activity in the somatosensory cortex was examined and showed a distinct disruption of the individual barrel structures, particularly within the anteriolateral barrel subfield (ALBSF) (Figure 7F). A reduction in overall area and number of barrels in *linc-Brn1b* mutants was also observed in the highly organized posteromedial barrel subfield (PMBSF). These findings, already consistent with the *Rorb* in situ hybridization data, were confirmed in P7 coronal sections immunostained with vGLUT2 (Figure 7G) and 5-HTT (Figure 7H), both of which specifically label the barrel structures. Analysis of both of these markers corroborates a specific impairment of the barrels within the ALBSF, and a general disorganization of individual barrel structures in the *linc-Brn1b*<sup>-/-</sup> mice.

Taken together, these results demonstrate the requirement of *linc-Brn1b* for the proper development of different classes of projection neurons within the cerebral cortex, and suggest that the loss of *linc-Brn1b* may potentially have broader implications for state-dependent cortical sensory processing.

## Discussion

In the post genomic era, thousands of long noncoding RNAs have been discovered as transcribed units in mammalian genomes. However, what fraction of these new transcripts have general functional significance in vivo is debated. While several studies have indicated a role for lincRNAs in diverse biological processes (Ponting et al., 2009; Rinn and Chang, 2012; Mercer and Mattick, 2013; Ulitsky and Bartel, 2013), it has been suggested that most transcripts could represent nonfunctional transcriptional by-products (Struhl, 2007; Kowalczyk et al., 2012). Early critical studies of knockout strains (e.g., *Xist* and *Tsix*) did find lncRNAs implicated in X inactivation to be required for life. Yet, of the relatively few lncRNA mouse models derived since, many have displayed subtle defects or no phenotype (Ripoche et al., 1997; Gordon et al., 2010; Anguera et al., 2011; Nakagawa et al., 2011; Zhang et al., 2012).

Other strategies using RNAi and xenografts to assess the function of lncRNAs in vivo have revealed interesting roles in development and tumor growth (Ulitsky et al., 2011; Wang et al., 2011; Yang et al., 2013). Together with difficulties in finding a phenotype in mouse models such as *Malat1*, *Neat1* (Nakagawa et al., 2011; Zhang et al., 2012), these findings have led some to suspect that acute inactivation of lncRNAs leads to stronger phenotypes than constitutive deletions, where compensatory events may occur. Therefore, our study, grounded in genetic deletions, demonstrates the important physiological insights that can be gleaned by constitutive lincRNA knockouts.

By leveraging our large-scale RNA-sequencing and genomics studies in combination with an integrated and multifaceted candidate selection pipeline, we were successful in finding physiologically essential lincRNAs. Through the initial characterization of the 18 lincRNAs knockout strains generated here, we found three that exhibit peri- or post-natal lethality and two additional ones with distinct developmental defects. Future studies will describe in detail other promising phenotypes.

A few recent studies have demonstrated that lincRNAs regulate neighboring protein coding genes and thereby may function as *cis*-enhancers or *cis*-regulatory elements (Ørom et al., 2010; Wang et al., 2011). Here, we were able to examine several lincRNAs and the regulation of the neighboring protein coding genes in a genetically defined context. We find one instance where the deletion of a lincRNA, *Fendrr*, phenocopies the neighboring protein coding gene despite a non-significant effect on *Foxf1* gene expression. Interestingly, chromatin predictions of enhancers (Shen et al., 2012) suggest this lincRNA may have enhancer-like properties (H3K4me1 and Pol II). However, our genetic analysis is not consistent with the notion of *Fendrr* as a chromatin signature-defined 'enhancer'. Specifically, Grote



et al observed a similar lethality phenotype in *Fendrr* mutants, yet their strategy (insertion of a strong transcription stop sequence) retained almost the entire DNA segment of *Fendrr*. This suggests that *Fendrr* cannot be acting simply as a DNA enhancer, although we cannot rule out the passive act of transcription as being functional. In our model, both the endogenous promoter and first exon remain. Moreover, the introduced *lacZ* reporter is actively transcribed as well, thus ruling out the passive act of transcription for *Fendrr* activity. Finally, since this lincRNA does not appear to act as a local enhancer and does not encode a small peptide based on PhyloCSF and Ribosome profiling analyses, we believe *Fendrr* to be a physiologically relevant and functional RNA molecule.

Interestingly, *Fendrr* mutant mice exhibit many features of *Foxf1a* protein coding gene heterozygous mutants (Mahlapuu et al., 2001). Since we do not observe a significant decrease in *Foxf1a* expression in *Fendrr* mutants, this raises the possibility that *Fendrr* could act downstream of *Foxf1a*. The FOXF1A protein plays an important role in the development of the lung and the gastrointestinal tract (Mahlapuu et al., 2001). Similarly, we found that *Fendrr* is expressed in the mesenchyme of these tissues and that homozygote mutant mice display lung and gastrointestinal tract defects most likely leading to perinatal lethality. Several cases of genetic deletions encompassing the neighboring protein coding gene *FOXF1*, and also *FENDRR* have been observed in neonates with the lethal lung disorder 'alveolar capillary dysplasia with misalignment of pulmonary veins' (ACD/MPV) (Szafranski et al., 2013). This supports our hypothesis that lincRNAs can be required for normal organ development and misregulated in pathological states, which could translate to human disease.

Deletion of both *Peril* and *Mdgt* also results in viability defects. *Mdgt* is transcribed from a bidirectional promoter shared with *Hoxd1*, which is located only 84 bp away from *Mdgt*. Although we do not exclude that *Hoxd1* expression could be affected by our *Mdgt* deletion strategy, the phenotype of homozygous *Hoxd1* mutants, which are viable and fertile, is in sharp contrast with *Mdgt*<sup>-/-</sup> mutants (Guo et al., 2010). Thus, deletion of *Mdgt* does not phenocopy *Hoxd1*, indicating a distinct function for this lincRNA. Interestingly, mice with a deletion of the adjacent *Hoxd3* protein coding gene display reduced viability similar to *Mdgt* (Condie and Capecchi, 1993). However, contrary to *Mdgt*<sup>-/-</sup>, the surviving *Hoxd3*<sup>-/-</sup> mice do not appear to have a growth defect, again suggesting a distinct function for *Mdgt*.

Neither the *Peril* nor *Mdgt* locus appears to be enriched for the signature enhancer modifications H3K4me1, H3K27ac in any public mouse ENCODE datasets. In the case of *Peril*, no genes were found significantly differentially expressed in knockout vs wild-type brains within ±1 Mb of *Peril* (Figure 3—figure supplement 2C; p<1.0, bootstrapped from 1000 random genomic intervals). Therefore, cis effects, as suggested by other enhancer-associated lincRNAs like *Hottip* (Wang et al., 2011), are unlikely. Collectively, these results suggest that lethality of the three lincRNAs observed in this study are likely not due to cis or enhancer-like RNA regulatory effects.

In addition to viability phenotypes, we also describe several developmental abnormalities such as body size and cortical defects in the *linc-Brn1b* mutant strain. Based on currently available public data, the *linc-Brn1b* locus does not appear to have enrichment for chromatin marks characteristic of enhancers, but rather the canonical H3K4me3/H3K36me3 signature of active transcription in the brain. In this case, however, deletion of *linc-Brn1b* did affect the expression of the neighboring protein coding gene. *linc-Brn1b* resides near *Brn1* (*Pou3f3*), a key transcription factor that shares redundant roles with the closely related paralog *Brn2* (*Pou3f2*) in upper layer cortical development. Interestingly, deletion of *linc-Brn1b* results in a significant reduction in the neighboring *BRN1* protein (Figure 6). However, in contrast to *linc-Brn1b* mutants, deletion of *Brn1* alone does not lead to defects in cortical lamination. Only when both *Brn1* and *Brn2* are deleted (*Brn1/2* double mutants) is a reduction in layers II–IV neurons observed (McEvelly, 2002; Sugitani et al., 2002; Dominguez et al., 2012). Despite the differential regulation of the adjacent *Brn1* gene, *linc-Brn1b* does not appear to act as a general cis-enhancer. The region has five other genes within ±1 Mb that are differentially expressed. However, this is not significant given the dramatic changes to the transcriptome in the *linc-Brn1b*<sup>-/-</sup> E13.5 telencephalon (Figure 3—figure supplement 2B; p<0.225, bootstrapped from 1000 random genomic intervals).

We do not observe a significant difference in expression of the paralogous *Brn2* protein coding gene in any *linc-Brn1b*<sup>-/-</sup> developmental stage analyzed when compared to wild type (Figure 6—figure supplement 1C). Thus, despite an observed decrease in *Brn1* expression in *linc-Brn1b*<sup>-/-</sup>, deletion of *linc-Brn1b* does not phenocopy deletion of the neighboring protein coding gene *Brn1*. In fact, ablation of *linc-Brn1b* results in a stronger phenotype than that observed for the neighboring protein



coding gene, suggesting additional roles for this lincRNA. In addition, we observe expression of *linc-Brn1b* in regions of the brain with no known expression of *Brn1*. Together, these observations suggest that lincRNAs adjacent to important developmental regulators may act on upstream and/or distinct pathways in addition to already reported *cis*-enhancer-like mechanisms (Ørom et al., 2010; Wang et al., 2011).

Further leveraging our genetically defined deletion we gleaned additional insights into lincRNA biology and transcriptional regulation in vivo. It has been noted that lincRNAs have a propensity to be transcribed from bidirectional promoters (Cabili et al., 2011; Pauli et al., 2012; Ulitsky and Bartel, 2013). New studies have also suggested that many lincRNAs transcribed from bidirectional promoters are unstable and likely non-functional transcripts (Almada et al., 2013; Sigova et al., 2013). Interestingly, *Fendrr* is transcribed from a bidirectional promoter. Here, we observe that deletion of *Fendrr* is lethal despite leaving an intact promoter, expression of the first exon and not interfering with Pol II dynamics at this locus. Similarly, RNA-Seq analysis of *linc-Brn1b* knockout revealed a strong and significant upregulation of *linc-Brn1a*, a lincRNA transcribed from a bidirectional promoter shared with *Brn1*. This incongruous activity on adjacent genes (decrease in *Brn1* protein coding gene and increase in *linc-Brn1a* transcribed from a bidirectional promoter with *Brn1*) observed in the *linc-Brn1b*<sup>-/-</sup> brain is suggestive of a functional role for the dynamically regulated *linc-Brn1a*, rather than transcriptional noise from bidirectional transcription from the protein coding gene promoter. Together, these findings confirm that not all lincRNAs transcribed from a bidirectional promoter are irrelevant transcriptional by-products, but rather suggests that some fraction of these transcripts play critical functional roles during development.

Our framework for lincRNA candidate selection for genetic analysis, based on RNA-sequencing catalogs and genomic studies, has led us to unexplored roles for lincRNAs in brain development. The brain (and the CNS more broadly) constitutes one of the most complex and fast evolving organs in the body. Here, it is likely that complex regulatory mechanisms of lincRNAs have evolved to build the layered control of gene expression necessary to generate the unparalleled cellular diversity and complex function of this organ. *linc-Brn1b* represents an example of a lincRNA that has an effect on development of the cerebral cortex. Beyond *linc-Brn1b*, many more lincRNAs from our screen have similar restricted patterns within progenitors in the VZ and SVZ of the telencephalon, possibly suggesting roles in neurogenesis. Others have highly cell-specific and dynamic in vivo expression patterns in distinct regions of the brain (Mercer et al., 2008). Thus, it is likely that more extensive work will reveal additional lincRNA mutants strains with additional brain and/or behavioral defects yet unexplored here. While rigorous behavioral studies will, in the future, determine whether loss of a lincRNA may result in specific behavioral abnormalities, we have made promising initial observations on two mutant strains. One mutant strain (*Spasm*) has tremors and a propensity to develop spastic movements upon handling, while a second strain (*linc-p21*) displays claspings of hind limbs when lifted by the tail (Sauvageau et al., unpublished).

As a whole, our 18 lincRNA knockout mouse models have revealed important aspects of lincRNA biology and constitute a useful resource for many future studies on the roles of lincRNAs in mammalian development, physiology and behavior.

## Materials and methods

### Mice

lincRNA knockout mice were generated by replacing the selected lincRNA gene with a *lacZ* cassette. Briefly, targeting constructs were constructed using VelociGene technology as described previously (Valenzuela et al., 2003). The VelociGene Allele Identification Numbers are shown in **Supplementary file 1A**. Linearized targeting constructs, generated by gap repair cloning containing mouse lincRNA upstream and downstream homology arms flanking a KOZAK-ATG-*lacZ*-pA-LoxP-hUb1-EM7-neo(superscript R)-pA-LoxP cassette, were electroporated into VGF1 hybrid mouse embryonic stem (ES) cells, derived from a 129S6S v/Ev female to a C57BL/6N male mating. Mouse ES cells carrying a heterozygous deletion of the lincRNA gene were identified by loss-of-function allele screening with 2 Taqman qPCR assays (**Supplementary file 1F**). Simultaneous replacement of the lincRNA gene with the *lacZ* cassette was confirmed by gain-of-allele Taqman assays against the *lacZ* and neomycin resistance cassette (**Supplementary file 1F**). Probes were labeled with 6-carboxy-fluorescein (FAM) on their 5' ends and BHQ-1 on their 3' ends. Targeted ES clones were introduced into an 8-cell stage mouse

embryo using the VelociMouse method (Poueymirou *et al.*, 2006). Mice were backcrossed once with C57BL/6J. Mutant mice were identified by genotyping for loss of lincRNA allele and gain of *lacZ* cassette. Toe clips, embryos or yolk sac were digested for 30 min at 95°C in 100 µl of 25 mM Sodium Hydroxide and 0.2 mM EDTA. Tissue digestion was neutralized by adding 100 µl of 40 mM Tris-HCl. PCR reactions using 4 µl of digested tissue with 10 mM *lacZ* specific and lincRNA gene specific primer pairs (Supplementary file 1C for sequences) were then performed and run on a 2% agarose gel. PCR conditions were as follows: 5 min at 95°C followed by 35 cycles of 30 s at 95°C, 45 s at 60°C, 30 s at 72°C and a final step at 72°C for 2 min. Mice were housed under controlled pathogen-free conditions (Harvard University's Biological Research Infrastructure) and experiments were approved by the Harvard University Committee on the Use of Animals in Research and Teaching. Viability of the 18 lincRNA mutant strains was determined at postnatal day 21 by genotyping the progeny of heterozygous intercrosses (Supplementary file 1C for genotyping primer sequences). In the case of lethal strains, the developmental stage at which lethality occurs was determined by genotyping of embryos at E14.5 and E18.5 and newborns. Respiratory function (*Fendrr* mutant strain) was evaluated in surgically delivered E18.5 embryos from heterozygous intercrosses (Eggan *et al.*, 2001). After cleaning of the airways, pups were placed on a 37°C warm pad and observed for sign of breathing.

### RNA isolation and mRNA-Seq library preparation

Total RNA from embryonic and postnatal mouse tissues, neural stem cells, and neurospheres was isolated using TRIzol (Life Technology, Carlsbad, CA)/chloroform extraction followed by spin-column purification (RNeasy mini kit, Qiagen, Venlo, Netherlands) according to the manufacturer instructions. RNA concentration and purity were determined using a Nanodrop (Thermo Fisher, Waltham, MA). RNA integrity was assessed on a Bioanalyzer (Agilent, Santa Clara, CA) using the RNA 6000 RNA chip. High-quality RNA samples (RNA Integrity Number  $\geq 8$ ) were used for library preparation. mRNA-seq libraries were constructed using the TruSeq RNA Sample Preparation Kit (Illumina, San Diego, CA) as previously described (Trapnell *et al.*, 2012a; Sun *et al.*, 2013). 500 ng total RNA was used as input for the TruSeq libraries from mouse tissues, and 200 ng for the libraries from neural stem cells and neurospheres. Prior to sequencing, libraries were run on a Bioanalyzer DNA7500 chip to assess purity, fragment size, and concentration. Libraries free of adapter dimers and with a peak region area (220–500 bp)  $\geq 80\%$  of the total area were sequenced. Individually barcoded samples were pooled and sequenced on the Illumina HiSeq 2000 platform.

### RNA-Seq analysis

Paired-end 101 bp reads were aligned to the mouse (mm9) reference genome assembly and, for the human neuronal differentiation time course also to the human (hg19) assembly, using Tophat2 (Trapnell *et al.*, 2009) with default options and assembled into transcripts with Cufflinks (Trapnell *et al.*, 2012a, 2012b). Aligned reads and assembled transcriptome catalog were used as input for Cuffdiff2 (Trapnell *et al.*, 2012a) to determine expression levels (FPKM, Fragments Per Kilobase per Million mapped reads) and differential expression between conditions using default options. CummeRbund v2.1 (<http://compbio.mit.edu/cummeRbund/>) was then used to process, index, and visualize the output of the Cuffdiff2 analyses. Gene set enrichment analysis (GSEA) and Guilt-by-Association analysis (GBA) were performed to predict the effect of gene expression changes on biological processes. Detailed description of GSEA and GBA analysis, RNA isolation, and libraries preparation are provided in the Extended Experimental Procedures. *Cis*-enhancer activity was tested by determining the number of genes with differential expression in a particular Knockout vs wild type contrast within  $\pm 1$  Mb window of the targeted lincRNA. 1000 random genomic intervals of the same size were obtained and interrogated in kind to determine how often the same number of differentially expressed (DE) genes could be identified. The ratio of intervals with DE genes  $\geq$  the number of DE genes in the target-flanking window to the number of iterations, provided a bootstrapped p value and false discovery rate estimate.

### Gene set enrichment analysis (GSEA) and Guilt-by-Association (GBA) analysis

GSEA was performed to predict the effect of the significant gene expression changes on biological processes in the knockout mice. For a given comparison of KO vs WT differential expression, all genes were rank-ordered by their Cuffdiff2 test-statistics and mouse gene identifiers were mapped to human

HUGO gene names. Genesets from the Reactome collection at mSigDB were obtained and for each gene set, the relative enrichment or depletion within our ranked list was determined via Mann-Whitney U-test. p values were corrected for multiple tests using the Benjamini-Hochberg method. Genesets with  $q < 0.001$  were selected and presented as a heatmap with color mapped to the Z-score of the Cuffdiff2 test statics for genes in the specific gene set relative to all genes. Redundant genesets were aggregated into higher-level biological processes.

Predictive GBA analysis for 17/18 tested lincRNAs was conducted as follows: Pearson correlation values of FPKM expression profiles were calculated for each lincRNA to all protein coding genes across a compendium of RNA-Seq samples (combination of in-house samples and samples from [Merkin et al., 2012]). Protein coding genes were then rank-ordered and subjected to the gene set enrichment analysis described above. Significant genesets for a given lincRNA represent the most likely pathways/biological processes for which this lincRNA may play a role.

## ***lacZ* expression and histology**

Expression of the knocked-in *lacZ* reporter gene was assessed in heterozygous mice. Embryos (from E13.5 to E18.5) were fixed in 4% paraformaldehyde (PFA) in phosphate buffered saline (PBS) overnight at 4°C prior to dissection of the brain, lung and respiratory tract, digestive tract, heart, and other organs. P7 brain, from *linc-Brn1b* mutant strain, were dissected from pups transcardially perfused with 4% paraformaldehyde (PFA), and fixed overnight at 4°C. The fixed tissues were rinsed three times at room temperature in PBS, 2 mM MgCl<sub>2</sub>, 0.01% deoxycholic acid, 0.02% NP-40. X-gal staining was performed by incubating the tissues for up to 16 hr at 37°C in the same buffer supplemented with 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide and 1 mg/ml X-gal. Staining reaction was stopped by washing three times in PBS at room temperature, followed by 2 hr post-fixation in 4% PFA at 4°C. Stained whole organs and sagittal brain sections were imaged using a Leica M216FA stereomicroscope (Leica Microsystems, Buffalo Grove, IL) equipped with a DFC300 FX digital imaging camera (Watson et al., 2008). Histology was performed at the Rodent Histopathology Service of the Dana Farber/Harvard Cancer Center Pathology Research Core. Embryos were harvested, fixed in Bouin's solution and embedded in paraffin. Microtome sections were stained with hematoxylin and/or eosin for histological analysis.

## **Nissl staining and immunocytochemistry**

Embryonic brains, dissected in cold PBS and fixed in 4% PFA/PBS overnight at 4°C, and P7 brains dissected from pups transcardially perfused with 4% PFA and post-fixed as described above, were processed for Nissl staining and immunofluorescence as previously described (Englund et al., 2005; Molyneaux et al., 2005). Nissl-stained and immunostained sections were imaged using a Nikon 90i fluorescence microscope equipped with a Retiga EXi camera (Q-IMAGING, Surrey, Canada) and acquired with Volocity image analysis software v4.0.1 (Perkin Elmer, Waltham, MA). For quantification of overall cortical thickness, cortical layers and number of CUX1<sup>+</sup>, CTIP2<sup>+</sup> and TLE4<sup>+</sup> cells within the primary somatosensory cortical area, anatomically matched sections were processed ( $n = 3$  *linc-Brn1b*<sup>-/-</sup>;  $n = 3$  wild-type, at P7). Boxes of 300 pixels in width and spanning the thickness of the cortex were superimposed at matched locations on each section, and the overall cortical thickness was measured as the distance from the pia to the white matter in each box, using ImageJ. Specific layer thicknesses were measured at the midpoint of the matched-location 300 pixel images for each of the TLE4<sup>+</sup>, CTIP2<sup>+</sup> and SATB2<sup>+</sup> immunofluorescence stainings using ImageJ. Layer VI thickness was measured as the distance between the dorsal edge of the TLE4<sup>+</sup> region and the white matter. Layer V thickness was determined by the span of the CTIP2<sup>+</sup> region, and layer II–IV thickness were measured as the SATB2<sup>+</sup> region between the dorsal edge of the CTIP2<sup>+</sup> stain and the pia. In each case results were expressed as mean  $\pm$  SEM. Cell counts of the specific neuronal subpopulations were obtained using the ITCN plugin for ImageJ and results were expressed as mean  $\pm$  SEM. A priori criteria were defined for analysis. Statistical analysis was performed using R unpaired Student's t test assuming equal variance was used for the pairwise comparisons.

## **Antibodies and image acquisition**

Primary antibodies and dilutions were as follows: anti- $\beta$ gal (CGAL-45; 1:500; Immunology Consultants Laboratory, Portland OR), anti-SATB2 (ab51502; 1:50; Abcam, Cambridge, UK), anti-CUX1 (sc-13024; 1:100; Santa Cruz Biotechnology, Dallas, TX), anti-pH3 (06-570; 1:500; Millipore, Billerica, MA), anti-TBR2

(1:2000; Gift from Robert F Hevner Lab), anti-CTIP2 (ab18465; 1:100; Abcam), anti-TLE4 (sc-9125; 1:100; Santa Cruz), anti-vGLUT2 (MAB5504; 1:50; Millipore), anti-5-HTT (PC177L; 1:1000; Millipore), anti-TUJ1 (mms-435P; 1:1000; Covance, Princeton, NJ). Secondary antibodies were from the Molecular Probes (Eugene, OR) Alexa series and were used at 1:750 dilution. Immunostained sections were imaged using a Nikon 90i fluorescence microscope equipped with a Retiga EXi camera (Q-IMAGING) and acquired with Volocity image analysis software v4.0.1 (Improvision).

## Cytochrome oxidase staining

*linc-Brn1b*<sup>-/-</sup> and wild-type littermates at P7 were anesthetized, perfused with 4% PFA, decapitated, and the brain rapidly removed. The brains were post-fixed in 4% PFA for 3 hr at 4°C. Cortices containing the barrel field were dissected and flattened as described (Welker and Woolsey, 1974), post-fixed in 4% PFA for 12–14 hr at 4°C and sectioned by using a vibratome (80 µm). Sections were incubated in phosphate buffer containing 0.5 mg/ml DAB, 0.18 mg cytochrome C, and 40 mg/ml sucrose for 3–5 hr at 37°C, rinsed, and mounted in Fluoromount-G (SouthernBiotech, Birmingham, AL) (Wong-Riley, 1979; Land and Simons, 1985).

## In situ hybridization

Nonradioactive in situ hybridization was performed on 40 µm vibratome sections mounted on superfrost slides (Fisher Scientific, Waltham, MA) as previously described (Berger and Hediger, 2001; Arlotta et al., 2005). The probe for *Svet1* was a gift from M Studer. Probes for *Rorb* (nt 1573-2087 of NM\_146095) and *Cux2* (nt 1069-11694 of NM\_007804) transcripts were generated by PCR from mouse brain cDNA and subcloned in pCRII-TOPO (Life Technologies, Carlsbad, CA). Antisense riboprobes were generated by *in vitro* transcription using SP6 or T7 polymerase (Roche Applied Science, Penzberg, Germany) as previously described. Sense probes were used as negative controls.

## Single molecule FISH

Single molecule FISH was performed as described by Raj et al. (2008). Briefly, oligonucleotide probes targeting and tiling *Peril* (48 probes) and *linc-Brn1b* (20 probes) were conjugated to Quasar 570 fluorophores and HPLC purified (Biosearch Technologies, Petaluma, CA). A list of the *Peril* and *linc-Brn1b* probes (sequence, position, and %CG content) is provided in **Supplementary file 1E**. Dissociated E14.5 cortical neurospheres or mouse ES cells were fixed in 2% formaldehyde for 10 min, washed twice with PBS, and permeabilized with 70% ethanol. The cells were then seeded onto previously gelatinized two-chamber cover glasses. Prior to hybridization, the cells were rehydrated in wash buffer containing 10% formamide and 2 × SSC for 5 min. Probes (0.5 ng/µl final concentration) were hybridized in 10% dextran sulfate, 10% formamide, and 2 × SSC at 37°C overnight. After hybridization, cells were washed twice with wash buffer at 37°C for 30 min (with DAPI added to the second wash for nuclear staining), and twice with 2 × SSC. After the SSC wash, the cells were equilibrated in anti-fade buffer (2 × SSC, 0.4% glucose, 10 mM Tris pH 8.0) for 3–5 min. Cells were mounted in 100 µl anti-fade buffer supplemented with 1 µl of glucose oxidase (G2133-10KU; Sigma-Aldrich, St. Louis, MO) and 1 µl of catalase (C3515-10 MG; Sigma-Aldrich) and immediately imaged with a LSM 700 Inverted Confocal microscope (Zeiss, Jena, Germany). 25 Z-stacks were taken per field, using DAPI and laser 639 for excitation.

## Human neural stem cell culture and differentiation

H1 human neural stem cells were prepared as described previously (Arnold et al., 2008; Goff et al., 2009) and grown at 37°C, 5% CO<sub>2</sub> on 1:4 diluted Matrigel-coated wells in neural proliferation medium (NPM; 50% DMEM/F12 Glutamax, 50% Neurobasal medium, 0.5X N2, 0.5X B27 without vitamin A, 20 ng/ml FGF [Life Technologies]). For differentiation, cells were plated at a density of 10<sup>6</sup> cells per well in a 6-well plate and allowed to proliferate for one day in the NPM medium. Neural induction was then initiated by withdrawal of FGF and addition of BDNF by switching the medium to neural differentiation medium (NDM; 100% Neurobasal medium, 1X B27 without vitamin A [Life Technologies], 10 ng/ml BDNF [Peprotech, Rocky Hill, NJ]). Differentiating cultures were maintained by refreshing NDM every other day until collection. Samples of these cultures were collected at days 0, 1, 2, and 4. Remaining cells (those designated for collection at days 5, 11, and 18) were replated at day 4 at a density 10<sup>6</sup> cells per well of a Poly-D-lysine/laminin-coated 6-well plate. Cells were harvested with Accutase (Stem Cell Technologies, Vancouver, Canada) and RNA collected as described above.



## Western blot

*linc-Brn1b*<sup>-/-</sup> and WT E14.5-derived neurospheres (passage 3) and E15.5 cortices were lysed in RIPA buffer (1% NP-40, 1% Na-deoxycholate, 0.2% SDS, 50 mM Tris 7.4, 500 mM NaCl) containing protease inhibitor cocktail (Roche Applied Science). Proteins were resolved by SDS-PAGE and electroblotted. Blots were probed sequentially for BRN1 (anti-BRN1, sc-6028-R, Santa Cruz) and GAPDH (anti-GAPDH, sc-365062-R, Santa Cruz) to control for protein loading. Immunoreactive bands were detected by enhanced chemiluminescence (SuperSignal, Thermo Fisher Scientific, Waltham, MA), and visualized with a Gel Doc (Biorad).

## Accession numbers

All primary RNA-Seq data are available at the Gene Expression Omnibus (GSE49581).

## Acknowledgements

We are grateful to Moran Cabili, Barbara Tazon-Vega and members of the Rinn Lab for discussions and critical comments on the manuscript and Olivia Plana for help with mouse colonies. We would like to thank Pamela Russel and Mitch Guttman for help with ribosome release score analyses, Roderick Bronson from the Rodent Histopathology Core at Dana-Farber/Harvard Cancer Center and Cathy MacGillivray from Harvard Stem Cell and Regenerative Biology Histology Core for help with the histopathology analyses. We also thank Biosearch Technologies for providing FISH probes, advice and expertise. JLR is supported by the NIH Directors New Innovator (DP2OD006670), P01 GM099117 and Center for Cell Circuits (P50 HG006193-01). This work was also supported by grants to PA from the NIH (NS062489, NS073124, NS078164), the New York Stem Cell Foundation, and the Harvard Stem Cell Institute. PA is a New York Stem Cell Foundation-Robertson Investigator. LAG. is a recipient of a NSF Postdoctoral Fellowship in Biology.

## Additional information

### Competing interests

VL: Employee of Regeneron Pharmaceuticals. GG: Employee of Regeneron Pharmaceuticals. GDY: Employee of Regeneron Pharmaceuticals. DF: Employee of Regeneron Pharmaceuticals. DMV: Employee of Regeneron Pharmaceuticals. The other authors declare that no competing interests exist.

### Funding

Funder	Grant reference number	Author
National Institutes of Health	DP2OD006670	John L Rinn
National Institutes of Health	P01 GM099117	John L Rinn
Center for Cell Circuits	P50 HG006193-01	John L Rinn
National Institutes of Health	NS062489	Paola Arlotta
National Institutes of Health	NS073124	Paola Arlotta
National Institutes of Health	NS078164	Paola Arlotta
New York Stem Cell Foundation		Paola Arlotta
Harvard Stem cell Institute		Paola Arlotta

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Author contributions

MS, LAG, SL, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article, Contributed unpublished essential data or reagents; BB, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; AFG, CG, DBS-G, EH, EL, MS, SCL, MM, MRS, MFD'E, JCM, Acquisition of data, Drafting or revising the article; WM, Analysis and interpretation of data, Drafting or revising the article; VL, GG, GDY, DF, MK, DMV, Drafting or revising the article, Contributed unpublished essential data or reagents; RPH, Analysis and interpretation of

data, Drafting or revising the article, Contributed unpublished essential data or reagents; PA, JLR, Conception and design, Analysis and interpretation of data, Drafting or revising the article, Contributed unpublished essential data or reagents

## Ethics

Animal experimentation: All animals were handled according to approved institutional animal care and use committee (IACUC) protocols (#11-13) of Harvard University.

## Additional files

### Supplementary files

• Supplementary file 1. (A) lincRNA genomic and targeted deletion coordinates (mm9). (B) Mass spectrometry peptide tags at background level identified for 4 lincRNA candidates (lif-over of lincRNA candidates to hg19). (C) Mouse genotyping primer sequences. (D) Quantitative RT-PCR primers. (E) Single molecule FISH probe sequences. (F) Taqman qPCR primers for targeted mES diagnostic.

DOI: [10.7554/eLife.01749.018](https://doi.org/10.7554/eLife.01749.018)

### Major datasets

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset ID and/or URL	Database, license, and accessibility information
Sauvageau M, Goff LA, Lodato S, Bonev B, Groff AF, Gerhardinger C, et al.	2013	Multiple knockout mouse models reveal lincRNAs are required for life and brain Development	GSE49581; <a href="http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE49581">http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE49581</a>	Publicly available at NCBI GEO ( <a href="http://www.ncbi.nlm.nih.gov/geo/">http://www.ncbi.nlm.nih.gov/geo/</a> ).

The following previously published datasets were used:

Author(s)	Year	Dataset title	Dataset ID and/or URL	Database, license, and accessibility information
Merkin J, Russell C, Chen P, Burge CB	2012	Evolutionary dynamics of gene and isoform regulation in mammalian tissues	GSE41637; <a href="http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE41637">http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE41637</a>	Publicly available at NCBI GEO ( <a href="http://www.ncbi.nlm.nih.gov/geo/">http://www.ncbi.nlm.nih.gov/geo/</a> ).
Ingolia NT, Lareau LF, Weissman JS	2011	Ribosome profiling of mouse embryonic stem cells reveals the complexity of mammalian proteomes	GSE30839; <a href="http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE30839">http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE30839</a>	Publicly available at NCBI GEO ( <a href="http://www.ncbi.nlm.nih.gov/geo/">http://www.ncbi.nlm.nih.gov/geo/</a> ).
Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, et al.	2012	A draft map of cis-regulatory sequences in the mouse genome	GSE29184; <a href="http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29184">http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29184</a>	Publicly available at NCBI GEO ( <a href="http://www.ncbi.nlm.nih.gov/geo/">http://www.ncbi.nlm.nih.gov/geo/</a> ).

## References

- Almada AE, Wu X, Kriz AJ, Burge CB, Sharp PA. 2013. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature* **499**:360–363. doi: [10.1038/nature12349](https://doi.org/10.1038/nature12349).
- Anguera MC, Ma W, Clift D, Namekawa S, Kelleher RJ, Lee JT. 2011. Tss produces a long noncoding RNA and has general functions in the germline, stem cells, and brain. *PLOS Genetics* **7**:e1002248. doi: [10.1371/journal.pgen.1002248](https://doi.org/10.1371/journal.pgen.1002248).
- Arlotta P, Molyneaux BJ, Chen J, Inoue J, Kominami R, Macklis JD. 2005. Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. *Neuron* **45**:207–221. doi: [10.1016/j.neuron.2004.12.036](https://doi.org/10.1016/j.neuron.2004.12.036).
- Arnold SJ, Huang G-J, Cheung AFP, Era T, Nishikawa S-I, Bikoff EK, Molnár Z, Robertson EJ, Groszer M. 2008. The T-box transcription factor Eomes/Tbr2 regulates neurogenesis in the cortical subventricular zone. *Genes and Development* **22**:2479–2484. doi: [10.1101/gad.475408](https://doi.org/10.1101/gad.475408).

- Berger UV**, Hediger MA. 2001. Differential distribution of the glutamate transporters GLT-1 and GLAST in tanyocytes of the third ventricle. *The Journal of Comparative Neurology* **433**:101–114. doi: [10.1002/cne.1128](https://doi.org/10.1002/cne.1128).
- Brunner AL**, Beck AH, Edris B, Sweeney RT, Zhu SX, Li R, Montgomery K, Varma S, Gilks T, Guo X, Foley JW, Witten DM, Giacomini CP, Flynn RA, Pollack JR, Tibshirani R, Chang HY, van de Rijn M, West RB. 2012. Transcriptional profiling of long non-coding RNAs and novel transcribed regions across a diverse panel of archived human cancers. *Genome Biology* **13**:R75. doi: [10.1186/gb-2012-13-8-r75](https://doi.org/10.1186/gb-2012-13-8-r75).
- Cabili MN**, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes and Development* **25**:1915–1927. doi: [10.1101/gad.17446611](https://doi.org/10.1101/gad.17446611).
- Chen B**, Schaevitz LR, McConnell SK. 2005. Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. *Proceedings of the National Academy of Sciences of the United States of America* **102**:17184–17189. doi: [10.1073/pnas.0508732102](https://doi.org/10.1073/pnas.0508732102).
- Condie BG**, Capecchi MR. 1993. Mice homozygous for a targeted disruption of Hoxd-3 (Hox-4.1) exhibit anterior transformations of the first and second cervical vertebrae, the atlas and the axis. *Development* **119**:579–595.
- Cubelos B**, Sebastian-Serrano A, Kim S, Moreno-Ortiz C, Redondo JM, Walsh CA, Nieto M. 2008. Cux-2 controls the proliferation of neuronal intermediate precursors of the cortical subventricular zone. *Cerebral Cortex* **18**:1758–1770. doi: [10.1093/cercor/bhm199](https://doi.org/10.1093/cercor/bhm199).
- Dominguez MH**, Ayoub AE, Rakic P. 2012. Pou-iii transcription factors (Brn1, Brn2, and Oct6) influence neurogenesis, molecular identity, and migratory destination of upper-layer cells of the cerebral cortex. *Cerebral Cortex* **23**:2632–2643. doi: [10.1093/cercor/bhs252](https://doi.org/10.1093/cercor/bhs252).
- Edwards AM**, Isserlin R, Bader GD, Frye SV, Willson TM, Yu FH. 2011. Too many roads not taken. *Nature* **470**:163–165. doi: [10.1038/470163a](https://doi.org/10.1038/470163a).
- Eggan KK**, Akutsu HH, Loring JJ, Jackson-Grusby LL, Klemm MM, Rideout WMW, Yanagimachi RR, Jaenisch RR. 2001. Hybrid vigor, fetal overgrowth, and viability of mice derived by nuclear cloning and tetraploid embryo complementation. *Proceedings of the National Academy of Sciences of the United States of America* **98**:6209–6214. doi: [10.1073/pnas.101118898](https://doi.org/10.1073/pnas.101118898).
- Englund C**, Fink A, Lau C, Pham D, Daza RAM, Bulfone A, Kowalczyk T, Hevner RF. 2005. Pax6, Tbr2, and Tbr1 are expressed sequentially by radial glia, intermediate progenitor cells, and postmitotic neurons in developing neocortex. *Journal of Neuroscience* **25**:247–251. doi: [10.1523/JNEUROSCI.2899-04.2005](https://doi.org/10.1523/JNEUROSCI.2899-04.2005).
- Gascoigne DK**, Cheetham SW, Cattenoz PB, Clark MB, Amaral PP, Taft RJ, Wilhelm D, Dinger ME, Mattick J. 2012. Pinstripe: a suite of programs for integrating transcriptomic and proteomic datasets identifies novel proteins and improves differentiation of protein-coding and non-coding genes. *Bioinformatics* **28**:3042–3050. doi: [10.1093/bioinformatics/bts582](https://doi.org/10.1093/bioinformatics/bts582).
- Goff LA**, Davila J, Swerdel MR, Moore JC, Cohen RI, Wu H, Sun YE, Hart RP. 2009. Ago2 Immunoprecipitation identifies predicted MicroRNAs in human embryonic stem cells and neural Precursors. *PLOS ONE* **4**:e7192. doi: [10.1371/journal.pone.0007192](https://doi.org/10.1371/journal.pone.0007192).
- Gomez JA**, Wapinski OL, Yang YW, Bureau J-F, Gopinath S, Monack DM, Chang HY, Brahic M, Kirkegaard K. 2013. The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon- $\gamma$  locus. *Behavioural Brain Research* **152**:743–754. doi: [10.1016/j.cell.2013.01.015](https://doi.org/10.1016/j.cell.2013.01.015).
- Gordon FE**, Nutt CL, Cheunsuchon P, Nakayama Y, Provencher KA, Rice KA, Zhou Y, Zhang X, Klibanski A. 2010. Increased expression of angiogenic genes in the brains of mouse Meg3-null embryos. *Endocrinology* **151**:2443–2452. doi: [10.1210/en.2009-1151](https://doi.org/10.1210/en.2009-1151).
- Grote P**, Witter L, Hendrix D, Koch F, Währisch S, Beisaw A, Macura K, Bläss G, Kellis M, Werber M, Herrmann BG. 2013. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Developmental Cell* **24**:206–214. doi: [10.1016/j.devcel.2012.12.012](https://doi.org/10.1016/j.devcel.2012.12.012).
- Guo T**, Mandai K, Condie BG, Wickramasinghe SR, Capecchi MR, Ginty DD. 2010. An evolving NGF-Hoxd1 signaling pathway mediates development of divergent neural circuits in vertebrates. *Nature Neuroscience* **14**:31–36. doi: [10.1038/nn.2710](https://doi.org/10.1038/nn.2710).
- Guttman M**, Rinn JL. 2012. Modular regulatory principles of large non-coding RNAs. *Nature* **482**:339–346. doi: [10.1038/nature10887](https://doi.org/10.1038/nature10887).
- Guttman M**, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL, Lander ES. 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**:223–227. doi: [10.1038/nature07672](https://doi.org/10.1038/nature07672).
- Guttman M**, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, Young G, Lucas AB, Ach R, Bruhn L, Yang X, Amit I, Meissner A, Regev A, Rinn JL, Root DE, Lander ES. 2011. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **477**:295–300. doi: [10.1038/nature10398](https://doi.org/10.1038/nature10398).
- Guttman M**, Garber M, Levin JZ, Donaghey J, Robinson J, Adiconis X, Fan L, Koziol MJ, Gnirke A, Nusbaum C, Rinn JL, Lander ES, Regev A. 2010. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nature Biotechnology* **28**:503–510. doi: [10.1038/nbt.1633](https://doi.org/10.1038/nbt.1633).
- Guttman M**, Russell P, Ingolia NT, Weissman JS, Lander ES. 2013. Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. *Behavioural Brain Research* **154**:240–251. doi: [10.1016/j.cell.2013.06.009](https://doi.org/10.1016/j.cell.2013.06.009).
- Jabaudon D**, Shnyder JS, Tischfield DJ, Galazo MJ, Macklis JD. 2012. ROR $\beta$  induces barrel-like neuronal clusters in the developing neocortex. *Cerebral Cortex* **22**:996–1006. doi: [10.1093/cercor/bhr182](https://doi.org/10.1093/cercor/bhr182).
- Khalil AM**, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES, Rinn JL. 2009. Many human large intergenic noncoding RNAs associate

- with chromatin-modifying complexes and affect gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **106**:11667–11672. doi: [10.1073/pnas.0904715106](https://doi.org/10.1073/pnas.0904715106).
- Koop KE, MacDonald LM, Lobe CG. 1996. Transcripts of Grg4, a murine groucho-related gene, are detected in adjacent tissues to other murine neurogenic gene homologues during embryonic development. *Mechanisms of Development* **59**:73–87. doi: [10.1016/0925-4773\(96\)00582-5](https://doi.org/10.1016/0925-4773(96)00582-5).
- Kowalczyk MS, Higgs DR, Gingeras TR. 2012. Molecular biology: RNA discrimination. *Nature* **482**:310–311. doi: [10.1038/482310a](https://doi.org/10.1038/482310a).
- Land PW, Simons DJ. 1985. Cytochrome oxidase staining in the rat Sml barrel cortex. *The Journal of Comparative Neurology* **238**:225–235. doi: [10.1002/cne.902380209](https://doi.org/10.1002/cne.902380209).
- Lin MF, Jungreis I, Kellis M. 2011. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics* **27**:i275–i282. doi: [10.1093/bioinformatics/btr209](https://doi.org/10.1093/bioinformatics/btr209).
- Mahlapuu M, Enerbäck S, Carlsson P. 2001. Haploinsufficiency of the forkhead gene Foxf1, a target for sonic hedgehog signaling, causes lung and foregut malformations. *Development* **128**:2397–2406.
- McEvilly RJ. 2002. Transcriptional regulation of cortical neuron Migration by POU domain factors. *Science* **295**:1528–1532. doi: [10.1126/science.1067132](https://doi.org/10.1126/science.1067132).
- Mercer TR, Dinger ME, Sunkin SM, Mehler MF, Mattick JS. 2008. Specific expression of long noncoding RNAs in the mouse brain. *Proceedings of the National Academy of Sciences of the United States of America* **105**:716–721. doi: [10.1073/pnas.0706729105](https://doi.org/10.1073/pnas.0706729105).
- Mercer TR, Mattick JS. 2013. Structure and function of long noncoding RNAs in epigenetic regulation. *Nature Structural and Molecular Biology* **20**:300–307. doi: [10.1038/nsmb.2480](https://doi.org/10.1038/nsmb.2480).
- Merkin J, Russell C, Chen P, Burge CB. 2012. Evolutionary dynamics of gene and isoform regulation in mammalian tissues. *Science* **338**:1593–1599. doi: [10.1126/science.1228186](https://doi.org/10.1126/science.1228186).
- Molyneaux BJ, Arlotta P, Hirata T, Hibi M, Macklis JD. 2005. Fezl is required for the birth and specification of corticospinal motor neurons. *Neuron* **47**:817–831. doi: [10.1016/j.neuron.2005.08.030](https://doi.org/10.1016/j.neuron.2005.08.030).
- Molyneaux BJ, Arlotta P, Menezes JRL, Macklis JD. 2007. Neuronal subtype specification in the cerebral cortex. *Nature Reviews Neuroscience* **8**:427–437. doi: [10.1038/nrn2151](https://doi.org/10.1038/nrn2151).
- Moseley ML, Zu T, Ikeda Y, Gao W, Mosemiller AK, Daughters RS, Chen G, Weatherspoon MR, Clark HB, Ebner TJ, Day JW, Ranum LP. 2006. Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nature Genetics* **38**:758–769. doi: [10.1038/ng1827](https://doi.org/10.1038/ng1827).
- Nakagawa S, Naganuma T, Shioi G, Hirose T. 2011. Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. *The Journal of Cell Biology* **193**:31–39. doi: [10.1083/jcb.201011110](https://doi.org/10.1083/jcb.201011110).
- Ørom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytznicki M, Notredame C, Huang Q, Guigo R, Shiekhattar R. 2010. Long noncoding RNAs with enhancer-like function in human cells. *Cell* **143**:46–58. doi: [10.1016/j.cell.2010.09.001](https://doi.org/10.1016/j.cell.2010.09.001).
- Pauli A, Valen E, Lin MF, Garber M, Vastenhouw NL, Levin JZ, Fan L, Sandelin A, Rinn JL, Regev A, Schier AF. 2012. Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Research* **22**:577–591. doi: [10.1101/gr.133009.111](https://doi.org/10.1101/gr.133009.111).
- Petersen CCH. 2007. The functional organization of the barrel cortex. *Neuron* **56**:339–355. doi: [10.1016/j.neuron.2007.09.017](https://doi.org/10.1016/j.neuron.2007.09.017).
- Ponting CP, Oliver PL, Reik W. 2009. Evolution and functions of long noncoding RNAs. *Behavioural Brain Research* **136**:629–641. doi: [10.1016/j.cell.2009.02.006](https://doi.org/10.1016/j.cell.2009.02.006).
- Poueymirou WT, Auerbach W, Frendewey D, Hickey JF, Escaravage JM, Esau L, Doré AT, Stevens S, Adams NC, Dominguez MG, Gale NW, Yancopoulos GD, DeChiara TM, Valenzuela DM. 2006. F0 generation mice fully derived from gene-targeted embryonic stem cells allowing immediate phenotypic analyses. *Nature Biotechnology* **25**:91–99. doi: [10.1038/nbt1263](https://doi.org/10.1038/nbt1263).
- Qureshi IA, Mattick JS, Mehler MF. 2010. Long non-coding RNAs in nervous system function and disease. *Brain Research* **1338**:20–35. doi: [10.1016/j.brainres.2010.03.110](https://doi.org/10.1016/j.brainres.2010.03.110).
- Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S. 2008. Imaging individual mRNA molecules using multiple singly labeled probes. *Nature Methods* **5**:877–879. doi: [10.1038/nmeth.1253](https://doi.org/10.1038/nmeth.1253).
- Ramos AD, Diaz A, Nellore A, Delgado RN, Park K-Y, Gonzales-Roybal G, Oldham MC, Song JS, Lim DA. 2013. Integration of genome-wide approaches identifies lncRNAs of adult neural stem cells and their progeny in vivo. *Cell Stem Cell* **12**:616–628. doi: [10.1016/j.stem.2013.03.003](https://doi.org/10.1016/j.stem.2013.03.003).
- Rinn JL, Chang HY. 2012. Genome regulation by long noncoding RNAs. *Annual Review of Biochemistry* **81**:145–166. doi: [10.1146/annurev-biochem-051410-092902](https://doi.org/10.1146/annurev-biochem-051410-092902).
- Ripoche MA, Kress C, Poirier F, Dandolo L. 1997. Deletion of the H19 transcription unit reveals the existence of a putative imprinting control element. *Genes and Development* **11**:1596–1604. doi: [10.1101/gad.11.12.1596](https://doi.org/10.1101/gad.11.12.1596).
- Sessa A, Mao C-A, Hadjantonakis A-K, Klein WH, Broccoli V. 2008. Tbr2 directs conversion of radial glia into basal precursors and guides neuronal amplification by indirect neurogenesis in the developing neocortex. *Neuron* **60**:56–69. doi: [10.1016/j.neuron.2008.09.028](https://doi.org/10.1016/j.neuron.2008.09.028).
- Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, Wagner U, Dixon J, Lee L, Lobanenkov VV, Ren B. 2012. A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**:116–120. doi: [10.1038/nature11243](https://doi.org/10.1038/nature11243).
- Sigova AA, Mullen AC, Molinier B, Gupta S, Orlando DA, Guenther MG, Almada AE, Lin C, Sharp PA, Giallourakis CC, Young RA. 2013. Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **110**:2876–2881. doi: [10.1073/pnas.1221904110](https://doi.org/10.1073/pnas.1221904110).



- Struhl K.** 2007. Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nature Structural and Molecular Biology* **14**:103–105. doi: [10.1038/nsmb0207-103](https://doi.org/10.1038/nsmb0207-103).
- Sugitani Y, Nakai S, Minowa O, Nishi M, Jishage K-I, Kawano H, Mori K, Ogawa M, Noda T.** 2002. Brn-1 and Brn-2 share crucial roles in the production and positioning of mouse neocortical neurons. *Genes and Development* **16**:1760–1765. doi: [10.1101/gad.978002](https://doi.org/10.1101/gad.978002).
- Sun L, Goff LA, Trapnell C, Alexander R, Lo KA, Hacısuleyman E, Sauvageau M, Tazon-Vega B, Kelley DR, Hendrickson DG, Yuan B, Kellis M, Lodish HF, Rinn JL.** 2013. Long noncoding RNAs regulate adipogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **110**:3387–3392. doi: [10.1073/pnas.1222643110](https://doi.org/10.1073/pnas.1222643110).
- Szafranski P, Dharmadhikari AV, Brosens E, Gurha P, Kolodziejska KE, Zhishuo O, Dittwald P, Majewski T, Mohan KN, Chen B, Person RE, Tibboel D, de Klein A, Pinner J, Chopra M, Malcolm G, Peters G, Arbuckle S, Guiang SF III, Hustead VA, Jessurun J, Hirsch R, Witte DP, Maystadt I, Sebire N, Fisher R, Langston C, Sen P, Stankiewicz P.** 2013. Small noncoding differentially methylated copy-number variants, including lncRNA genes, cause a lethal lung developmental disorder. *Genome Research* **23**:23–33. doi: [10.1101/gr.141887.112](https://doi.org/10.1101/gr.141887.112).
- Tarabykin V, Stoykova A, Usman N, Gruss P.** 2001. Cortical upper layer neurons derive from the subventricular zone as indicated by Svet1 gene expression. *Development* **128**:1983–1993.
- Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L.** 2012a. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature Biotechnology* **31**:46–53. doi: [10.1038/nbt.2450](https://doi.org/10.1038/nbt.2450).
- Trapnell C, Pachter L, Salzberg SL.** 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**:1105–1111. doi: [10.1093/bioinformatics/btp120](https://doi.org/10.1093/bioinformatics/btp120).
- Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L.** 2012b. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**:562–578. doi: [10.1038/nprot.2012.016](https://doi.org/10.1038/nprot.2012.016).
- Tsai M-C, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY.** 2010. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**:689–693. doi: [10.1126/science.1192002](https://doi.org/10.1126/science.1192002).
- Ulitsky I, Bartel DP.** 2013. lincRNAs: genomics, evolution, and mechanisms. *Cell* **154**:26–46. doi: [10.1016/j.cell.2013.06.020](https://doi.org/10.1016/j.cell.2013.06.020).
- Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP.** 2011. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147**:1537–1550. doi: [10.1016/j.cell.2011.11.055](https://doi.org/10.1016/j.cell.2011.11.055).
- Valenzuela DM, Murphy AJ, Frendewey D, Gale NW, Economides AN, Auerbach W, Poueymirou WT, Adams NC, Rojas J, Yasenchak J, Chernomorsky R, Boucher M, Elsasser AL, Esau L, Zheng J, Griffiths JA, Wang X, Su H, Xue Y, Dominguez MG, Noguera I, Torres R, Macdonald LE, Stewart AF, DeChiara TM, Yancopoulos GD.** 2003. High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. *Nature Biotechnology* **21**:652–659. doi: [10.1038/nbt822](https://doi.org/10.1038/nbt822).
- Visel A, Zhu Y, May D, Afzal V, Gong E, Attanasio C, Blow MJ, Cohen JC, Rubin EM, Pennacchio LA.** 2010. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature* **464**:409–412. doi: [10.1038/nature08801](https://doi.org/10.1038/nature08801).
- Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, Lajoie BR, Protacio A, Flynn RA, Gupta RA, Wysocka J, Lei M, Dekker J, Helms JA, Chang HY.** 2011. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**:120–124. doi: [10.1038/nature09819](https://doi.org/10.1038/nature09819).
- Watson CM, Trainor PA, Radziejewicz T, Pelka GJ, Zhou SX, Parameswaran M, Quinlan GA, Gordon M, Sturm K, Tam PPL.** 2008. Application of lacZ transgenic mice to cell lineage studies. *Methods in Molecular Biology* **461**:149–164. doi: [10.1007/978-1-60327-483-8\\_10](https://doi.org/10.1007/978-1-60327-483-8_10).
- Welker C, Woolsey TA.** 1974. Structure of layer IV in the somatosensory neocortex of the rat: description and comparison with the mouse. *The Journal of Comparative Neurology* **158**:437–453. doi: [10.1002/cne.901580405](https://doi.org/10.1002/cne.901580405).
- White JK, Gerdin A-K, Karp NA, Ryder E, Buljan M, Bussell JN, Salisbury J, Clare S, Ingham NJ, Podrini C, Houghton R, Estabel J, Bottomley JR, Melvin DG, Sunter D, Adams NC, Sanger Institute Mouse Genetics Project, Tannahill D, Logan DW, Macarthur DG, Flint J, Mahajan VB, Tsang SH, Smyth I, Watt FM, Skarnes WC, Dougan G, Adams DJ, Ramirez-Solis R, Bradley A, Steel KP.** 2013. Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. *Behavioural Brain Research* **154**:452–464. doi: [10.1016/j.cell.2013.06.022](https://doi.org/10.1016/j.cell.2013.06.022).
- Wong-Riley M.** 1979. Changes in the visual system of monocularly sutured or enucleated cats demonstrable with cytochrome oxidase histochemistry. *Brain Research* **171**:11–28. doi: [10.1016/0006-8993\(79\)90728-5](https://doi.org/10.1016/0006-8993(79)90728-5).
- Yang L, Lin C, Jin C, Yang JC, Tanasa B, Li W, Merkurjev D, Ohgi KA, Meng D, Zhang J, Evans CP, Rosenfeld MG.** 2013. lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. *Nature* **500**:598–602. doi: [10.1038/nature12451](https://doi.org/10.1038/nature12451).
- Yoon J-H, Abdelmohsen K, Srikantan S, Yang X, Martindale JL, De S, Huarte M, Zhan M, Becker KG, Gorospe M.** 2012. lncRNA-p21 suppresses target mRNA translation. *Molecular Cell* **47**:648–655. doi: [10.1016/j.molcel.2012.06.027](https://doi.org/10.1016/j.molcel.2012.06.027).
- Zhang B, Arun G, Mao YS, Lazar Z, Hung G, Bhattacharjee G, Xiao X, Booth CJ, Wu J, Zhang C, Spector DL.** 2012. The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Reports* **2**:111–123. doi: [10.1016/j.celrep.2012.06.003](https://doi.org/10.1016/j.celrep.2012.06.003).
- Zhu J, Sanborn JZ, Diekhans M, Lowe CB, Pringle TH, Haussler D.** 2007. Comparative genomics search for losses of long-established genes on the human lineage. *PLOS Computational Biology* **3**:e247–e247. doi: [10.1371/journal.pcbi.0030247](https://doi.org/10.1371/journal.pcbi.0030247).
- Zimmer C.** 2004. Dynamics of Cux2 expression suggests that an early pool of SVZ precursors is fated to become upper cortical layer neurons. *Cerebral Cortex* **14**:1408–1420. doi: [10.1093/cercor/bhh102](https://doi.org/10.1093/cercor/bhh102).